Rubin Wang
Fanji Gu
Editors

Advances in Cognitive Neurodynamics (II)

Proceedings of the Second International Conference on Cognitive Neurodynamics – 2009

Springer

Advances in Cognitive Neurodynamics (II)

Rubin Wang · Fanji Gu
Editors

Advances in Cognitive Neurodynamics (II)

Proceedings of the Second International Conference on Cognitive Neurodynamics - 2009

Springer

Editors
Prof. Rubin Wang
East China University of Science
and Technology
Institute for Cognitive Neurodynamics
130 Meilong Road
200237 Shanghai
People's Republic of China

Prof. Fanji Gu
East China University of Science
and Technology
Institute for Cognitive Neurodynamics
130 Meilong Road
200237 Shanghai
People's Republic of China

ISBN 978-90-481-9694-4 e-ISBN 978-90-481-9695-1
DOI 10.1007/978-90-481-9695-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2008928127

Preface

The 1st International Conference on Cognitive Neurodynamics (ICCN'07) was held in Shanghai in 2007. About 300 hundred experts from more than 30 countries and areas attended that conference and made a great success. Thus, in the editorial board meeting of the journal *Cognitive Neurodynamics* held during the conference, we decided to develop it to be a series conference holding every two years. Now the second one was just held in the beautiful city of Hangzhou, China, November 15–19, 2009.

As what we claimed in the editorial of the journal *Cognitive Neurodynamics*, cognitive neurodynamics is a fast growing new scientific front combining dynamic analysis with cognitive science and neuroscience experiments, which spans the entire range of brain cognition from ionic channels to whole brain activities, and which is combined by the conceptions of nonlinear neurodynamics operating simultaneously at and across all levels. The developments in these two years seem to verify our statement. Many important progresses have been made, and the 2nd conference offered us a wonderful opportunity to review the progress and discuss its problems.

As what we expected, the 2nd conference was also successful as the 1st one. There were 8 plenary talks given by the leading scientists in this field and 20 oral sessions and one poster session. The topics of the contributions cover almost all the branches of cognitive neurodynamics, from micro-, meso- to macroscopic cognitive neurodyanmics, their applications and some related topics, especially including neural coding, realistic neural networks, oscillation and synchronization, neural population dynamics, sensory and motor dynamics, EEG, fMRI and brain imaging, global cognitive function, multiscalar neurodynamics, neural computing, brain computer interface, and cognition disorder. It's also our great pleasure to notice the high quality of the contributions.

We are glad that we have a wonderful cooperation with Tamagawa Dynamic Brain Forums (DBF) organized by the Japanese "Gang of Five" on nonlinear dynamics. They gave generous support to this conference. They and RIKEN BSI organized a symposium for the conference and invited many speakers to give wonderful talks. These two series meetings focus almost on the same field, but with different scales and depth of discussion, now they support each other. We do believe

that these meetings will promote the progress in this field. And the 3rd conference will be organized by them in Sapporo, Japan in June, 2011. The papers invited by DBF are published as the second section of this volume. We also would like to express our gratitude to Profs. Francois-Benoit Vialatte, Justin Dauwels, and Jianting Cao for organizing a mini-symposium on Neural dynamics of brain disorders; Profs. Liqing Zhang, Tomasz M. Rutkowski, Toshihisa Tanaka, Jianting Cao for organizing one on Emerging Technologies for Brain Computer Interfaces; Profs. Qishao Lu, Jianzhong Su, Jianxue Xu for organizing one on Firing Oscillations and Patterns in Neuronal Networks; and Profs. Hans Liljenström and Hans Braun for organizing one on Multiscalar Neurodynamics – from Physiology to Systems Theory.

Taking this opportunity, we would also like to thank all the sponsors, especially our university – East China University of Science & Technology, the host university of ICCN'09 – Zhejiang University, our journal "Cognitive Neurodynamics" published by Springer, The National Natural Science Foundation of China, Brain Science Center of Tamagawa University, Chinese Society for Neuroscience, Chinese Society of Theoretical and Applied Mechanics, IEEE Singapore Computational Intelligence Chapter, International Neural Network Society, Japanese Neural Network Society, RIKEN Brain Science Institute, and many other universities in China.

Shanghai
1st December, 2009

Rubin Wang
Fanji Gu

Sponsors and Organizers

This conference is sponsored by

East China University of Science and Technology (ECUST), China
Zhejiang University, China
"Cognitive Neurodynamics" published by Springer Science+Business Media B.V., The Netherlands

Co-Sponsored by

Brain Science Center, Tamagawa University, Japan
Brain Science Research Center, KAIST, Korea
Chinese Society for Neuroscience (CSN)
Chinese Society of Theoretical and Applied Mechanics (CSTPM)
International Neural Network Society (INNS)
Japanese Neural Network Society (JNNS)
Research Center for Brain Science, Fudan University, China
RIKEN Brain Science Institute (RIKEN BSI), Japan
Shanghai Society for Nonlinear Science (SSNS)

Organized by

East China University of Science and Technology (ECUST)
Zhejiang University, China

Supported by

The National Natural Science Foundation of China (NSFC)
Beijing Institute of Technology, China
Beijing University of Aeronautics and Astronautics, China
Beijing University of Technology, China
Shanghai Jiaotong University, China
Shanghai University, China
Shanxi Normal University, China
South East University, China
Tongji University, China
Xi'an Jiaotong University, China

Organizing Committee

Honorary Chairs:

Amari, Shun-ichi (RIKEN Brain Science Institute, Japan)
Freeman, Walter J. (University of California, Berkeley, USA)
Grossberg, Stephen (Boston University, USA)
Haken, Hermann (University of Stuttgart, Germany)
Richmond, Barry J. (NIMH/NIH/DHHS, USA)
Tang, Xiaowei (Zhejiang University, China)
Taylor, John (King's College, UK)

General Chair:

Wang, Rubin (East China University of Science and Technology, China)

Co-Chair:

Li, Guang (Zhejiang University, China)
Hu, Haiyan (Beijing Institute of Technology, China)
Ishikawa, Masumi (Kyushu Institute of Technology, Japan)
Lu, Qishao (Beijing University Aeronautics and Astronautics, China)
Sandner, G.uy (University Louis Pasteur, France)
Tsukada, Minoru (Tamagawa University, Japan)

International Advisory Committee Chairs:

Adeli, Hojjat (The Ohio State University, USA)
Chu, Jian (Zhejiang University, China)
Kötter, Rolf (Radboud University Nijmegen, The Netherlands)
Schöner, Gregor (Ruhr-Universität Bochum, Germany)
Tsuda, Ichiro (Hokkaido University, Japan)
Wang, Deliang (The Ohio State University, USA)
Usui, Shiro (RIKEN BSI, Japan)

Members:

Aihara, Kazuyuki (The University of Tokyo, Japan)
Cao, Jinde (Southeast University, China)

Cateau, Hide (New York University, USA)
Chen, Liqun (Shanghai University, China)
Deco, Gustavo (Universitat Pomeu, Fabra, Spain)
Érdi, Peter (Kalamazoo College, KFKI Research Institute for Particle and Nuclear of the Hungarian Academy of Sciences, Hungary)
Fukai, Tomoki (Brain Science Institute, RIKEN, Japan)
Hayashi, Hatsuo (Kyushu Institute of Technology, Japan)
Hoffmann, Klaus-Peter (Ruhr University Bochum, Germany)
Kamel, Mohamed (University of Waterloo, Canada)
Kaski, Samuel (Helsinki University of Technology, Finland)
König, Peter (University Osnabrück, Germany)
Li, Chaoyi (Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, P. R. China)
Longtin, André (University of Ottawa, Canada)
Maass, Wolfgang (Technische Universität Graz, Austria)
Narayanan, Srini (University of California, Berkeley, USA)
Olds, James L. (George Mason University, USA)
Principe, Jose C. (University of Florida, USA)
Shou, Tiande (Fudan University, China)
Sporns, Olaf (Indiana University, USA)
Sun, Mingui (University of Pittsburgh, USA)
Szeto, K.Y. (Hong Kong University of Science and Technology, Hong Kong)
Teich, Malvin Carl (Boston Univeristy, USA)
Tirri, Henry (University of Helsinki, Helsinki University of Technology, NRC Nokia Research Center, Finland)
Trappenberg, Thomas (Dalhousie University, Canada)
van Leeuwen, Cees (RIKEN Brain Science Institute, Japan)
Ventriglia, Francesco (Institute of Cybernetics of CNR, Italy)
Yamguchi, Yoko (RIKEN Brain Science Institute (BSI), Japan)
Yang, Fusheng (Tsinghua University, China)
Zhang, Wei (Beijing University of Technology, China)
Zhu, Weiqiu (Zhejiang University, P.R. China)

Program Committee Chair:

Gu, Fanji (Shanghai Society for Biophysics/Fudan University, China)
Egan, Gary (University of Melbourne, Australia)
Kozma, Robert (The University of Memphis, USA)
Lee, Soo-Young (BSRC, KAIST, Korea)
Liljenström, Hans (Swedish University of Agricultural Sciences, Sweden)
Nara, Shigetoshi (Okayama University, Japan)
Shi, Hongbo (East China University of Science and Technology, China)
Wang, Lipo (Nanyang Technological University, Singapore)

Members:

Aertsen, Ad (Albert-Ludwigs Universitat Freiburg, Germany)
Andras, Peter (University of Newcastle, UK)
beim Graben, Peter (Reading University, UK)
Braun, Hans (Philipps University of Marburg, Germany)
Brette, Romain (Ecole Normale Supérieure, Paris, France)
Cao, Jianting (Sitama Institute of Technology, Japan)
Del Moral Hernandez, Emilio (University of Sao Paulo, Brazil)
Depireux, Didier (University of Maryland, USA)
Ding, Mingzhou (University of Florida, USA)
Erwin, Harry (University of Sunderland, UK)
Feldman, Jerry (ICSI, University of California at Berkeley, USA)
Franceschini, Nicolas (CNRS/University de la Méditerranée, France)
Fujii, Hiroshi (Kyoto Sangyo University Japan)
Gao, Shangkai (Tsinghua University, China)
Giese, Martin (University Clinic Tübingen, Germany)
Goldfarb, Lev (University of New Brunswick, Canada)
Graham, Lyle J. (Laboratory of Neurophysics and Physiology, CNRS, France)
Gu, Xingsheng (East China University of Science and Technology, P.R. China)
Hellgren Kotaleski, Jeanette (Royal Institute of Technology, Sweden)
Hofmann, Ulrich G. (University of Lübeck, Germany)
Jiang, Tianzi (Institute of Automation, Chinese Academy of Sciences, P.R. China)
Kashimori, Yoshiki (University of Electro-Communications, Japan)
Kay, Leslie M. (University of Chicago, USA)
Kendall, Graham (University of Nottingham, UK)
Lappe, Markus (Westfälische Wilhelms University, Germany)
Lee, Vincent C S (Monash University, Australia)
Liang, Hualou (University of Texas, USA)
Liang, Peiji (Shanghai Jiaotong University China)
Matsuoka, Kiyotoshi (Kyushu Institute of Technology, Japan)
Neskovic, Pedja (Brown University, USA)
Ogmen, Haluk (University of Houston, USA)
Safronov, Boris V. (IBMC, Portugal)
Tong, Qinye (Zhejiang University, China)
van Ooyen, Arjen (Vrije Universiteit, The Netherlands)
van Pelt, Jaap (Netherlands Institute for Neurosciences, The Netherlands)
Wright, James (Auckland University, New Zealand)
Wu, Si (Institute of Neuroscience, CAS, China)
Xu, Jian (Tongji University, China)
Xu, Jianxue (Xi'an Jiaotong University, P.R. China)
Young, Eric D. (Johns Hopkins University, USA)

Zhang, Liming (Fudan University, China)
Zhang, Liqing (Shanghai Jiaotong University, China)
Zhang, Yi (University of Electronic Science and Technology of China, China)
Zhou, Jing (Shanghai University, China)

Secretary General:

Wang, Ying (East China University of Science and Technology, China)
Wang, You (Zhejiang University, China)

Webmaster

Zhang, Jin (Zhejiang University, China)

Local Organization Committee

Chairperson:

Li, Guang (Zhejiang University)

Members:

Tong, Qinye (Zhejiang University, China)
Hu, Meng (Zhejiang University, China)
Zheng, Junbao (Zhejiang Sci-Tech University, China)

Contents

Part I
Plenary Talk

Dynamics of Learning In Hierarchical Models – Singularity and Milnor Attractor

Shun-ichi Amari, Tomoko Ozeki, Florent Cousseau, and Haikun Wei

Abstract We study the dynamics of learning in a hierarchical model such as multilayer perceptron. Such a model includes singularities, which affects its dynamics seriously. The Milnor type attractors appear, because of the singularity. We will show its trajectories explicitly, and present the topological nature of the singularities.

Keywords Dynamics of learning · On-line learning · Multilayer-perceptron · Singularity · Milnor attractor

1 Introduction

Neural learning takes place in the parameter space of a model network. A model has often a hierarchical structure of which typical examples are the multilayer perceptron and the Gaussian mixture radial basis network. Since the behavior of a network is subject to noise and training examples are given randomly, the dynamics of learning is a stochastic process, and a trajectory of learning fluctuates randomly.

The parameter space of a hierarchical network has singularities due to the permutation invariance of elements of hidden layers. The behavior of networks becomes the same in some region of the parameter space. Such regions form singularities of the parameter space and a network loses identifiability at singularities. The metrical structure of the space is given by the Fisher information matrix, but it degenerates at singularities, because of the non-identifiability. Singularities in a statistical model and the dynamical behaviors of learning have been studied in a number of papers [1–6]. We summarize these results, and give a new idea on the topological aspects related to singularities.

There arise lots of strange behaviors when a statistical model is singular [3]. The Cramer-Rao paradigm of estimation does not holds, and Bayesian priors become

S.-i. Amari (✉)
RIKEN Brain Science Institute, Saitama 351-0198, Japan
e-mail: amari@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_1,

singular. Singularities have strong effects on the behaviors of dynamics of learning. The natural gradient learning method [2] is known to have good performances even at singularities. The present paper studies dynamics of learning near singularities. We give explicit solutions of trajectories of the average learning equations in a neighborhood of singularities. This proves that singularities become the Milnor type attractors [7], and this is the reason why learning behaviors are retarded, called the plateau phenomena.

We further study the topological structure of singularities, and give a new coordinate system by the blow-down technique of algebraic geometry. Such a coordinate system is useful for studying the behavior of the dynamics of natural gradient learning.

2 Dynamics of On-Line Learning

Let us assume that the input-output behavior of a network is given by $y = f(\boldsymbol{x}, \boldsymbol{\theta}) + \varepsilon$, where $\boldsymbol{x} = (x_1, \cdots, x_n)$ is an input vector, y is a scalar output, ε is a standard Gaussian noise subject to $N(0, 1)$, and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)$ is parameters which are modifiable. In our case of supervised learning, the network modifies its parameter $\boldsymbol{\theta}$ by receiving a series of input examples $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_t$ and corresponding outputs $y_1, \cdots, y_t$ which a teacher gives,

$$y_i = f_0(\boldsymbol{x}_i) + \varepsilon,\ i = 1, \cdots, t. \tag{1}$$

The standard gradient learning algorithm is written as

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \nabla l(\boldsymbol{x}_t, y_t, \boldsymbol{\theta}_t), \tag{2}$$

where

$$l(\boldsymbol{x}, y, \boldsymbol{\theta}) = \frac{1}{2}\{y - f(\boldsymbol{x}, \boldsymbol{\theta})\}^2 \tag{3}$$

is the squared error function, $\nabla = \partial/\partial\boldsymbol{\theta}$ and η is a learning constant. This is a stochastic difference equation, because $\boldsymbol{x}_i$ is randomly generated and y_i includes noise. By introducing the error function

$$e(\boldsymbol{x}, y, \boldsymbol{\theta}) = f(\boldsymbol{x}, \boldsymbol{\theta}) - f_0(\boldsymbol{x}) - \varepsilon, \tag{4}$$

and using the continuous time version neglecting the fluctuation term, we have the average learning equation,

$$\frac{d}{dt}\boldsymbol{\theta}(t) = \dot{\boldsymbol{\theta}}_t = \eta \langle e(\boldsymbol{x}, y, \boldsymbol{\theta}_t) \nabla f(\boldsymbol{x}, \boldsymbol{\theta}_t) \rangle. \tag{5}$$

The actual trajectory fluctuates around the average one. The fluctuations do not accumulate, when the dynamics is stable.

3 Singularities of Hierarchical Networks

We consider a multilayer perceptron and a simple Gaussian mixture network. A multilayer perceptron has the input–output behavior,

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_{i=1}^{k} w_i \varphi\left(J_i \cdot \boldsymbol{x}\right) + \varepsilon. \tag{6}$$

where φ is a sigmoid function.

The parameter of a perceptron is specified by parameters $\boldsymbol{\theta} = (J_1, w_1, \cdots, J_k, w_k)$. A network does no have identifiability for some regions of the parameter space S in the above hierarchical networks. We explain this by two typical examples:

(1) When $w_i = 0$, the term $w_i \varphi\left(J_i \cdot \boldsymbol{\pi}\right)$ vanishes. Hence J_i does not affect the behavior $f(\boldsymbol{x}, \boldsymbol{\theta})$ and is not identifiable on the subspace $w_i = 0$.
(2) When $J_i = J_j$, the sum of two terms, $w_i \varphi\left(J_i \cdot \boldsymbol{x}\right) + w_j \varphi\left(J_j \cdot \boldsymbol{x}\right)$ depends only on $\left(w_i + w_j\right)$. Hence, each of w_i and w_j is not identifiable on the subspace $J_i = J_j$.

Hence, some parameters are not identifiable on the singular region

$$\mathcal{R}_{ij} = \left\{\boldsymbol{\theta} \,\middle|\, w_i w_j \left(J_i - J_j\right) = 0\right\}. \tag{7}$$

4 Critical Singular Region

We consider a typical situation of two hidden neurons, $k = 2$. Only one neuron suffices in the singular region $\mathcal{R} = \mathcal{R}_{12}$, where $J_1 = J_2$ or one of w_1 and w_2 is 0. In this case, its input–output function is written by using a single neuron,

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = w \varphi\left(J \cdot \boldsymbol{x}\right). \tag{8}$$

We introduce a new coordinate system $\boldsymbol{\xi} = (\boldsymbol{v}, w, \boldsymbol{u}, z)$, which uses

$$\boldsymbol{u} = J_2 - J_1, \tag{9}$$

$$z = \frac{w_1 - w_2}{w_1 + w_2}, \tag{10}$$

and the other parameters $\boldsymbol{v}$ and w are fixed. Then, the singular region is represented by $\boldsymbol{u} = 0$ or $z = \pm 1$. A singular region $\mathcal{R}_{12}$ includes three parts, one given by $\boldsymbol{u} = 0$, that is, $J_1 = J_2$, which we call the overlapping singular region, and the other is $z = \pm 1$, that is $w_1 = 0$ or $w_2 = 0$, in which one neuron is eliminated. We have $R_{12} = R_1 \cup R_2$,

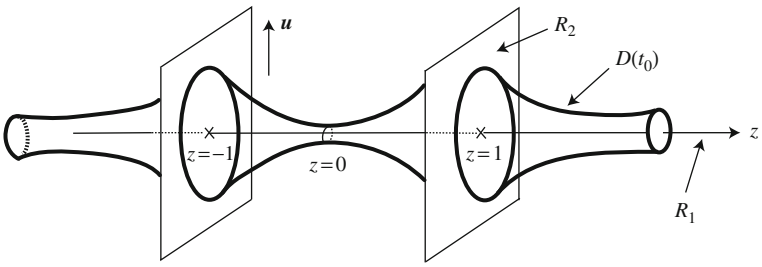


Fig. 1 Singular region and its neighborhood

$$R_1(\boldsymbol{v}, w) = \{\boldsymbol{\xi} \,|\boldsymbol{u} = 0, z : \textit{arbitrary}; \boldsymbol{v} \textit{ and } w \textit{ fixed}\}, \tag{11}$$

$$R_2(\boldsymbol{v}, w) = \{\boldsymbol{\xi} \,|\boldsymbol{u} : \textit{arbitrary}; z = \pm 1, \boldsymbol{v} \textit{ and } w \textit{ fixed}\}. \tag{12}$$

In the subspace in which $\boldsymbol{v}$ and w are fixed constants, R_1 is a line in which z changes, and R_2 consists of two n-dimensional surfaces. The singular region R_{12} is a composite of one line and two surfaces, see Fig. 1.

5 Dynamical Flow Near Singularities

We study the dynamical flow near the singularity $R_1(\boldsymbol{v}, w)$, by using the Taylor expansion with respect to $\boldsymbol{u}$. From this, we have the average learning equation is written as

$$\dot{\boldsymbol{u}} = \eta \frac{w}{2}(1 - z^2)\left\langle e\varphi''(\boldsymbol{v} \cdot \boldsymbol{x})(\boldsymbol{u} \cdot \boldsymbol{x})\boldsymbol{x}\right\rangle + O(|\boldsymbol{u}|^2), \tag{13}$$

$$\dot{z} = -\frac{1 + z^2}{w}\eta\left\langle \varphi'(\boldsymbol{v} \cdot \boldsymbol{x})(\boldsymbol{u} \cdot \boldsymbol{x})\right\rangle + O(|\boldsymbol{u}|^2). \tag{14}$$

We can integrate (13) and (14) in S^* where $\boldsymbol{v}$ and w take the optimal values, giving the flows or trajectories of dynamics in the neighborhood of $\boldsymbol{u} = 0$ of S^*. They are given by

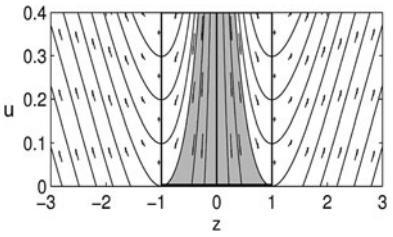


Fig. 2 Trajectories

$$\frac{1}{2}\boldsymbol{u}^T \cdot \boldsymbol{u} = \frac{2w^{*2}}{3} \log\left[\frac{(z^2+3)^2}{|z|}\right] + c \tag{15}$$

where c is an arbitrary constant. See Fig. 2 for the trajectories.

6 Milnor Attractor

The line R_1 $(\boldsymbol{v}^*, w^*)$ in S^* is a critical line of the dynamics of learning. Under a certain condition we show that the line is divided into two parts such that one part is stable and the other is unstable. In such a case, there are trajectories attracted to the stable part, and the basin of attraction has a finite measure. This is a Milnor type attractor, different from a saddle. More strongly the basin of attraction includes a neighborhood of this part, implying all the points near this part are once attracted to it. After being attracted to it, the parameters still move randomly in R_1 $(\boldsymbol{v}^*, w^*)$ by stochastic fluctuations until they reach the unstable part. Then, the parameters leave R_1 eventually.

We can prove that R_1^* is unstable (saddle) when $A = \left(A_{ij}\right)$, $A_{ij} = \langle e\phi_{ij}\rangle$, includes both positive and negative eigenvalues. When w^*A is positive-semi-definite, the interval $1 < z^2$ is attractive, and the remaining part $1 > z^2$ is repulsive. Hence, the interval $1 < z^2$ behaves like a Milnor attractor. This is the plateau phenomenon, taking long time before getting rid of it. When w^*A is negative-semi-definite, the part $1 > z^2$ is attractive and the other part is repulsive.

7 Topology of Singularity

We have shown that the singular region R consists of the three parts. One is the line specified by $\boldsymbol{u} = 0$, and the others are two submanifolds given by $z = 1$ and $z = -1$ (see Fig. 1). All the functions $f(\boldsymbol{x}, \boldsymbol{\xi})$ in $R(\boldsymbol{v}, w)$ are equal, having a common one behavioral function.

In order to see the topological structure of the singularity, we fix a pair $(\boldsymbol{v}, w)$ and consider the subspace $S(\boldsymbol{v}, w)$ where $\boldsymbol{v}$ and w are fixed and $\boldsymbol{u}$ and z are free. We map it to the space of input-output functions $\mathcal{F} = \{f(\boldsymbol{x})\}$. Here, all the points in R are mapped to one point given by (8). By this mapping, the singularity region $\mathcal{R}(\boldsymbol{v}, w)$ is reduced to one point and is "resolved" except for one point. This is the blowing-down of singularity.

In order to demonstrate the image of $S(\boldsymbol{v}, w)$ in $\mathcal{F}$, let us introduce a new parameter t, which represents how far a function $f(\boldsymbol{x}, \boldsymbol{\xi})$ is from the singular region $R(\boldsymbol{v}, w)$ given by (8). We introduce a new parameter

$$t = \left(1 - z^2\right) u^2, \tag{16}$$

which is 0 at the singularity R, and represents the Riemannian distance $\sqrt{g_{11}}t$ between R and $f(\boldsymbol{x}, \boldsymbol{\xi})$ when t is small, where g_{11} is the Fisher information in the direction of t.

Let us consider the set of points in S which are located within distance less than t_0 from R,

$$D(t_0) = \left\{ \boldsymbol{\xi} | \left(1 - z^2\right) u^2 \leq t_0 \right\}. \tag{17}$$

This is a level surface (Fig. 1), where t_0 denotes the distance form the singularity. When $t_0 = 0$, $D(0) = R$ which is the union of one line $\boldsymbol{u} = 0$ and two subspaces $z = \pm 1$. Likewise, $D(t_0)$ consists of three parts; one in the region $|z| < 1$ and the others in the regions $z > 1$ and $z < -1$. In each part, the boundary $\partial D(t_0)$ of $D(t_0)$ is a cylinder, which is obtained by rotating a curve $\left(1 - z^2\right) u^2 = t_0$ along the z-axis.

Let us visualize the image of the cylinder $D(t_0)$ in the middle part $|z| < 1$. The boundary $\partial D(t_0)$ is a cylinder having two coordinates $e \in S^n$ and $z \in [-1, 1]$. The two open sides of the cylinder, that is, the left and right sides corresponding to $z \to 1$ and $z \to -1$ approach the apex as $t \to 0$, touch R.

8 Conclusions

The dynamics of learning suffers from the plateau phenomenon ubiquitously, and it has gradually been understood that this is caused by the symmetry of its structure. The symmetry is caused by the permutation group of hidden units, and it gives rise to non-identifiability of the parameters in the region on which the permutation group acts invariantly. The Fisher information degenerates on such a region, and the residue class divided by the group has topological singularities.

We have elucidated the structure of the singularities. Moreover, we gave detailed mathematical analysis of the dynamics of learning near singularities by using a simple model, and obtained the trajectories explicitly. We proved that the singularity forms a Milnor attractor, which is not a stable equilibrium but has a finite measure of the basin of attraction. Hence a trajectory of learning has once attracted in the line, fluctuating around the line until it reaches unstable region, and eventually gets rid of it. This is the plateau phenomenon.

References

1. Cousseau, F., Ozeki, T., Amari, S.: Dynamics of learning in multilayer perceptrons near singularities. IEEE. Trans. Neural. Netw. **19** (2008) 1313–1328.
2. Amari, S.: Natural gradient works efficiently in learning. Neural. Comput. **10** (1998) 251–276.
3. Amari, S., Park, H., Ozeki, T.: Singularities affect dynamics of learning in neuromanifolds. Neural. Comput. **18** (2006) 1007–1065.
4. Fukumizu, K., Amari, S.: Local minima and plateaus in hierarchical structures of multilayer perceptrons. Neural. Netw. **13** (2000) 317–327.

5. Wei, H., Zhang, J., Cousseau, F., Ozeki, T., Amari, S.: Dynamics of learning near singularities in layered networks. Neural. Comput. **20** (2008) 813–843.
6. Wei, H., Amari, S.: Dynamics of learning near singularities in radial basis function networks. Neural. Netw. **21** (2008) 989–1005.
7. Milnor, J.: On the Concept of Attractor. Commun. Math. Phys. **99** (1985) 177–195.

Chaotic Dynamics, Episodic Memory, and Self-identity

Ichiro Tsuda

Abstract The hippocampus has been considered responsible for the formation of episodic memory. It has also been pointed out that the hippocampus plays an important role in imagination, which is related to future events. The fact that atrophy of the hippocampus could lead to Alzheimer's disease implies that the network structure of the hippocampus may provide fields for the creation of internal time corresponding to the past, present, and future. We present a hypothesis that the hippocampus plays a role in the formation of self-identity via interactions with the lateral prefrontal cortex.

Keywords Chaotic dynamics · Episodic memory · Self-identity · Internal time

1 Introduction

Memory and inference processes form the core cognitive functions. The neural mechanisms that underlie these functions have recently been highlighted. In particular, studies have focused on neural correlates of dynamic memory such as episodic memory. Tulving defined episodic memory as a memory that involves information acquired from individual experience [1]. Thus, remembering could depend on individual recollections, as we apprehend a fraction of the episode that we actually experienced, according to our ability to attend selectively. Many experimental facts have shown that the hippocampal networks are responsible for the formation of episodic memory. Here, we would like to highlight the specific facts that we think are most closely related to the contents of this article.

Fact 1. Damage to the hippocampus causes both retrograde and anterograde amnesia [2, 3].

I. Tsuda (✉)
Research Institute for Electronic Science, Hokkaido University, Sapporo 001-0020, Hokkaido, Japan
e-mail: tsuda@math.sci.hokudai.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_2,

Fact 2. Place cells observed in the rat hippocampus respond sensitively to the actual spatial locations occupied by the rat [4].

Fact 3. Atrophy of the hippocampus may cause Alzheimer's disease [5].

In addition, clinical studies of patients with a hippocampal lesion revealed that damage to the hippocampus can cause the inability to imagine new things [6].

Fact 4. The hippocampus is responsible for imagining new (future) experiences [6].

As this process is akin to composing a story, it is closely related to the formation of episodic memories. Furthermore, Maguire et al. stated that "episodic memory and imagining or constructing events share striking similarities such as imagery, sense of presence, retrieval of semantic information and multimodal details, and narrative structure". This finding is crucial for constructing a theory for self-identity, in which past, present and future events are connected.

In this article, we briefly review our recent study on information representation of episodic events in the model hippocampus, and show that the prediction of Cantor coding was verified by experiments using rat hippocampal slices. We also present a new theory of self-identity. Finally, we propose a hypothesis that the hippocampus forms self-identity via the interactions with the lateral prefrontal cortex.

2 Chaotic Itinerancy in CA3

The overall CA3 network possesses massively recurrent connections among excitatory pyramidal cells, and also inhibitory cells, which may locally inhibit the excitatory cells. These inhibitory cells may be inhibited by inhibitory cells in a different area of the brain; i.e., the septum. According to experimental observations [7], the septum's inhibitory cells are excited almost periodically, and this excitation is synchronized with the θ-rhythms, which is a brain wave with a 5–8 Hz frequency range. Thus, these inhibitory cells inhibit the inhibitory cells in almost periodic fashion.

Since Marr's pioneering work [8], many researchers have proposed hypotheses on the operation of episodic memory in terms of associative memory [9, 10]. However, an association dynamics with the same initial pattern input simply shows a relaxation process to an attractor, which may represent a memory associated with the input pattern. Thus, classical theory does not account for a sequence of episodic events. This can also be justified by the fact that the CA3 metric space concerned defines a measure of the similarity of patterns themselves but not of a sequence of patterns. We have proposed another hypothesis on the operation of episodic memory that extends the classical theory of associative memory to capture the dynamic and successive association of memory [11, 12]; this hypothesis is based on the dynamic associative memory models [13–15].

A key to the dynamics of memory is suggested by the following. In the phase during which the disinhibition signal from the septum is on, the main dynamics of CA3 arises from the recurrent network of pyramidal cells that is responsible for a

single association in memory. Then, the input signal triggers attractor dynamics, whereby a pattern associated with the input pattern is formed. Such an attractor pattern becomes a representation of the memory associated with that input. During this phase, a memory may be represented by a geometric attractor. On the other hand, in the phase during which the disinhibition signal from the septum is off, the direct inhibition of the recurrent network produces instability in the dynamics of memory. A geometric attractor becomes a quasi-attractor [11, 16], which is associated with the appearance of unstable direction; however, most dynamical orbits are attracted to this type of quasi-attractor. This instability generates a dynamic link among quasi-attractors, so that successive associations in memory can be produced. During this phase, a quasi-attractor is an attractor-ruin of the original attractor.

One can also describe this type of transitory dynamics in terms of chaotic itinerancy [17–19], as the transition process can be described by chaotic dynamics, which is high dimensional in some cases and low dimensional in others. We extracted a transition law in the latter phase of disinhibitory inputs, according to the correlations of memory patterns defined by the critical circle map that possesses fixed point attractors in a Milnor's sense [20]. For a detailed description of the mechanism of chaotic itinerancy, see reference [21]. Recently, both aperiodic and noisy transitions between several definite states of neural activity of the hippocampal CA3 were observed in the laboratory [22]. The computational results of our model CA3 revealed that CA3 may represent the dynamic association of memories via chaotic itinerancy and that episodic events may be represented by the chaotic itinerancy that occurs in CA3 as episodic memories.

3 Iterated Function Systems and Cantor Coding in CA1

CA1 networks include fewer recurrent connections than CA3 networks; instead they include mainly local negative feedback connections via inhibitory cells. The septum's inhibitory cells also inhibit the inhibitory cells of CA1, which are synchronized with the θ-rhythm. The connections between CA3 and CA1 within the hippocampus are almost unidirectional. The CA1 is considered a contracting system by tuning the control parameters, which can be derived from the absence of recurrent connections, and the presence of negative feedback connections by inhibitory neurons. Thus, in this case, the whole hippocampus can be viewed as a chaos-driven contracting system.

We investigated a biology-oriented model that represents the physiological neural networks of CA1. To make a single neuron model, we used the two-compartment model proposed by Pinsky and Rinzel [23]. In the CA1 model [24], an input to CA1 is a temporal sequence of spatial patterns, which is considered as an output of CA3. We investigated the dynamic behavior of membrane potentials produced by the model CA1 neurons. We found finite subsets of Cantor sets in both subthreshold and superthreshold dynamics. The resulting set has several elements, each of which represents finite fractions of the input temporal sequence. Because the input temporal

sequence is chaotic (i.e., ordered but aperiodic), as its length increases, the temporal sequence contains a great variety of finite sample paths starting from different initial patterns, because the input temporal sequence is chaotic, that is, ordered but aperiodic. Thus, when the length becomes infinite, the set becomes a mathematical Cantor set. Furthermore, the distribution of membrane potentials in the model CA1 neurons follows a bimodal distribution whose minimum corresponds to the single-neuron threshold. This result suggests the possibility of decoding the information embedded in Cantor sets via a pulse train output from pyramidal cells. However, we do not possess a concrete method for extracting it, therefore, the decoding problem still remains unsolved.

We also investigated the "rule dynamics" that creates the Cantor sets and found a set of contracting affine transformations. We found a set of contacting affine transformations. This finding suggests that a similar rule dynamics emerges to the one in Iterated Function System (IFS). Therefore, one can predict that the CA1 network creates a set of contracting affine transformations, by which Cantor sets are produced, where Cantor sets encode input temporal sequences [24].

We further developed an experimental system using rat hippocampal slices. For the input temporal sequences of patterns, we found a set of contracting affine transformations in the return maps of membrane potentials of CA1 neurons [25]. We also found a hierarchically classified pattern of membrane potentials, which seems to be yielded by IFS dynamics [26]. Here, a hierarchy of membrane potentials means a hierarchy of compression of the temporal sequence. Therefore, our theoretical predictions were demonstrated experimentally. The CA1 network transforms temporal sequences into spatial patterns, where the similarity between temporal sequences can be measured by a Hausdorff metric. In contrast, in CA3, as stated previously, we can only measure the similarity between patterns.

4 Rhythm and Simultaneity

In this section, we discuss the simultaneity in human perception and cognition, based on the works of Ernst Peppel [27]. This subject is crucial for a theory that connects the oscillatory behaviors of neural activity with self-identity. Peppel defined a "time window of simultaneity" as the input time width of sensory stimuli of the same kind during which we receive these stimuli as a single sensation. For instance, an increase in the input time width between successive click sounds given to the right and left ears (where the duration of one click is 1 ms) allows the discrimination of these sounds as two inputs provided at an interval greater than 4 ms. This kind of time window of simultaneity is around 10 ms for touch, and around 20–30 ms for vision. Thus, the time window varies according to the type of sensory stimulus. It should be noted that this time window corresponds to the time width that allows the perception of two successive stimuli, and differs from the time width necessary to discriminate the specific sequence of the stimuli. The latter was estimated by Peppel et al. as 30–40 ms, which applies to all kinds of sensory stimuli. As this time width

corresponds to the time window necessary for identification of each stimulus, here we call it a "time window of identity".

Another time width necessary to identify the kind of sensory stimulus given varies from 240 to 340 ms for auditory or visual perception, during which the response frequency becomes a local maximum at each 30–40 ms interval. From these observations, Peppel proposed the following hypothesis.

Hypothesis 1 (Peppel). Relaxation oscillations of ∼30 Hz in the brain synchronize each other and function as the neural mechanism that underlies the judgment of sensory stimuli.

These frequencies of oscillation corresponds to a range of slow γ-waves, which correlate with attention, judgment, and determination of behavior [28]. This fact suggests that γ-waves are the basis for the judgment of sensory stimuli. This kind of function is a prerequisite for the formation of episodic memory.

5 Logic for Self-identity

What is self-identity? Identifying a matter A means differentiating A from a distinct B matter. We conceived an inference process that may underlie this act. A possible inference process is as follows: "A is not B other than A". According to classical logic, the statement "A is not B other than A" is equivalent to "A is not $\neg A$", where the symbol $\neg$ indicates negation. Thus, the statement "A is not B other than A" is equivalent to the statement "A is A", which can represent simultaneity in a time domain. In other words, self-identity can be replaced by simultaneity in a time domain, within the framework of classical logic. On the other hand, in our context, self-identity should differ from simultaneity. In order to justify this difference, we explicitly introduced a *unit time* to logic, whereby the inference process underlying logic can be represented in terms of dynamical systems [29–31]. Here, we termed this inference process a *step inference*. A step inference consists of two processes: the logical process from premise to consequence, and the substitution process of the obtained consequence with the premise; in either or both processes, a unit time is assumed. Here, a variable is a truth value of the statement, which takes 0 or 1; 1 for "true" and 0 for "false". It can be extended to take a real value in the unit interval [0, 1]. For a while, we assume dichotomy.

Let us write down the equations of motion for simultaneity and self-identity. The equation of motion for simultaneity is derived from the successive inference of the statement "A is A". The equations of motion for self-identity are derived from the successive inference of the statement "A is not B, and B is not A". Let A and B be the truth values for the statements A and B, respectively. The subscript n denotes a discrete time that represents the succession of inference.

Simultaneity

$$A_{n+1} = A_n \tag{1}$$

Self-identity

$$A_{n+1} = 1 - B_n \tag{2}$$

$$B_{n+1} = 1 - A_n \tag{3}$$

or, equivalently;

$$A_{n+1} = A_{n-1} \tag{4}$$

Thus, simultaneity is a special case of self-identity in the sense that Eq. (4) contains Eq. (1). Moreover, the solution of the dynamical system for simultaneity is only fixed points, while the solution of the dynamical system for self-identity consists of fixed points, i.e., $(A, B) = (1, 0)$ and $(0, 1)$, and a period-two periodic orbit, i.e., $(1, 1) \leftrightarrow (0, 0)$, the latter of which indicates recurrence of "true" and "false". This means the following. Finding a fixed point of inference is sufficient for recognition of a matter as itself (i.e., simultaneity). On the other hand, two inference processes, such as deriving "false" from "true" and then deriving "true" from "false" are necessary for the recognition of a matter as itself, while discriminating from others (i.e., self-identity).

We now proceed to a case of continuous time in a real axis and a continuous variable representing a continuous truth value defined in a unit interval [0, 1]. Let the derived discrete dynamical systems, i.e., maps, be viewed as an Euler's discretization of the underlying differential equation. We assume $\Delta t = 1$.

From Eq. (1), the supposed original differential equation for simultaneity will be as follows.

$$\frac{dA}{dt} = 0 \tag{5}$$

The solution is constant. This means that "true" is derived from the initial condition "true" or that "false" is derived from the initial condition "false".

In contrast, we obtained similar differential equations for self-identity.

$$\frac{d(A + B)}{dt} = 1 - (A + B) \tag{6}$$

The asymptotic solution of $A + B$ is unity. This means that the consequence from classical logic $A = 1$ and $B = 0$ or $A = 0$ and $B = 1$ holds. This result coincides with the result derived from the differentiation of Eq. (4). Therefore, we obtain the fixed points that are equivalent to the logical consequence derived using classical logic. In other words, the truth values in classical logic are recovered in this asymptotic limit of a step inference. These differential equations do not yield any periodic orbits. It is known that classical logic does not contain "time". From our present consideration, this characteristic of classical logic stems from the dynamical behavior in an infinite time limit in an infinitesimal time interval. In this limit, one cannot distinguish between "simultaneity" and "self-identity". However, one can discriminate "self-identity" from "simultaneity" by introducing a finite width of time during which one can infer the truth of statements.

This logical inference of differential type, which is equivalent to an inference in classical logic, may be processed in the lateral prefrontal cortex, based on the input sensory stimulations [32]. According to Maguire's hypothesis, the hippocampus is responsible for integrating the past, present and future. On the other hand, in our present theory, we state that the formation of self-identity is necessary for the sense of the present, which can be a basis of a sense of presence, where a finite time window is prerequisite, which can be around 30 ms. Thus, the information contents that are transferred from the lateral prefrontal cortex to the hippocampus will include information on the step inference. In conclusion, we propose the following hypothesis.

Hypothesis 2. A neural process used to correctly identify sensory stimuli requires a finite time window of approximately 30 ms and uses an internal step-inference process embedded in the interactive neural systems of the hippocampus, sensory cortices, and lateral prefrontal cortex.

Acknowledgments The author would like to express his special thanks to Minoru Tsukada, Yasunori Fukushima, Shigeru Kuroda, and Yutaka Yamaguti for their invaluable discussions on both theoretical and experimental studies of the hippocampal function. This work was partially supported by Grant-in-Aid for Scientific Research – on Priority Areas "Integrative Brain Research" (18019002), partially supported by Grant-in-Aid for Scientific Research on Priority Areas "Understanding of Mobiligence" (18047001), partially supported by Grant-in-Aid for Scientific Research (B) (18340021), and partially supported by Grant-in-Aid for Exploratory Research (17650056) – all from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. Tulving, E.: Episodic and semantic memory. In Tulving, E., Donaldson, W., eds.: Organization of Memory. New York: Academic Press (1972) 381–403.
2. Scoville, W.B., Milner, B.: Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatr. **20** (1957) 11–21.
3. Zola-Morgan, S., Squire, L.R., Amaral, D.G.: Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field CA1 of hippocampus. J. Neurosci. **6** (1986) 2950–2967.
4. O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map: preliminary evidence from unit activity in the freely moving rat. Brain Res. **34** (1971) 171–175.
5. Dalla Barba, G., Nedjam, Z., Dubois, B.: Confabulation, executive functions, and source memory in Alzheimer's disease. Cogn. Neuropsychol. **16** (1999) 385–398.
6. Hassabis, D., Kumaran, D., Vann, S.D., Maguire, E.A.: Patients with hippocampal amnesia cannot imagine new experiences. Proc. Natl. Acad. Sci. USA. **104** (2007) 1726–1731.
7. Buzsaki, G., Functions for interneuronal nets in the hippocampus. Can. J. Physiol. Pharmacol. **75** (1997) 508–515.
8. Marr, D., Simple memory: a theory for archicortex. Phil. Trans. R. Soc. Lond. B. **262** (1971) 23–81.
9. Treves, A., Rolls, E.T.: Computational analysis of the hippocampus in memory. Hippocampus. **4** (1994) 374–391.
10. McClelland, J.L., McNaughton, B.L., O'Reilly, R.C.: Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. **102** (1995) 419–457.

11. Tsuda, I.: Towards an interpretation of dynamic neural activity in terms of chaotic dynamical systems. Behav. Brain Sci. **24** (2001) 793–847.
12. Tsuda, I., Kuroda, S.: Cantor coding in the hippocampus. Jpn. J. Ind. Appl. Math. **18** (2001) 249–258.
13. Tsuda, I.: Dynamic link of memories — chaotic memory map in nonequilibrium neural networks. Neural Netw. **5** (1992) 313–326.
14. Nara, S., Davis, P.: Chaotic wandering and search in a cycle-memory neural network. Prog. Theor. Phys. **88** (1992) 845–855.
15. Adachi, M., Aihara, K.: Associative dynamics in a chaotic neural network. Neural Netw. **10** (1997) 83–98.
16. Haken, H.: Beyond attractor neural networks for pattern recognition. Nonlinear Phenomena Complex Syst. **9** (2006) 163–172.
17. Ikeda, K., Otsuka, K., Matsumoto, K.: Maxwell-Bloch turbulence. Prog. Theor. Phys. **99**(Suppl.) (1989) 295–324.
18. Kaneko, K.: Clustering, coding, switching, hierarchical ordering, and control in network of chaotic elements. Physica D. **41** (1990) 137–172.
19. Tsuda, I.: Chaotic itinerancy as a dynamical basis of hermeneutics of brain and mind. World Futures. **32** (1991) 167–185.
20. Milnor, J.: On the concept of attractor. Commun. Math. Phys. **99** (1985) 177–195.
21. Tsuda, I.: Hypotheses on the functional roles of chaotic transitory dynamics. Chaos. **19** (2009) 015113-1–015113-10.
22. Sasaki, T., Matsuki, N., Ikegaya, Y.: Metastability of active CA3 networks. J. Neurosci. **17** (2007) 517–528.
23. Pinsky, P.F., Rinzel, J.: Intrinsic and network rhythmogenesis in a reduced traub model for CA3 neurons. J. Comput. Neurosci. **1** (1994) 39–60.
24. Yamaguti, Y., Kuroda, S., Tsuda, I.: A mathematical model for Cantor coding in the hippocampus. in preparation (2009).
25. Kuroda, S., Fukushima, Y., Yamaguti, Y., Tsukada, M., Tsuda, I. Iterated function systems in the hippocampal CA1. To be published in Cogn. Neurodyn. (2009).
26. Fukushima, Y., Tsukada, M., Tsuda, I., Yamaguti, Y., Kuroda, S.: Cantor-like coding in hippocampal CA1 pyramidal neurons. Cogn. Neurodyn. **1** (2007) 305–316.
27. Peppel, E.: Grenzen des Bewußtseins: Über Wirklichkeit und Welterfahrung (Stuttgart: Deutsche Verlags-Anstalt GmbH, 1985); Ishiki no Naka no Jikan (Translated into Japanese by T. Tayama and K. Ogata, Tokyo: Iwanami-Publ. 1995: Japanese translation is from original German version); Mindworks: Time and Conscious Experience (Translated into English by T. Artin, H. Brace, New York: Jovanovich Publishers, 1988).
28. Freeman, W.J., How Brains Make Up Their Minds (London: 1999).
29. Grim, P. Self-reference and chaos in fuzzy logic. IEEE Trans. Fuzzy Syst. **1** (1993) 237–253.
30. Mar, G., Grim, P.: Pattern and chaos: new images in the semantics of paradox. Nous. **25** (1991) 659–693.
31. Tsuda, I., Tadaki, K.: A logic-based dynamical theory for a genesis of biological threshold. BioSystems. **42** (1997) 45–64.
32. Pan, X., Sawa, K., Tsuda, I., Tsukada, M., Sakagami, M.: Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. Nat. Neurosci. **11** (2008) 703–712.

Stochastic Modeling of Neuronal Responses

Barry J. Richmond

Abstract The sequence of spikes in a train is considered to be a neuronal code. We have shown that a simple model in which the spikes are stochastically thrown down with the probabilities governed by a rate function over time estimated from the spike density over time. The numbers of spikes is determined by measuring the number of spikes over the interval of interest, usually several hundred milliseconds. By analyzing data from this model, which is instantiated by order statistics, neuronal spike trains can be decoded instant-by-instant as they unfold over time.

Keywords Neural code · Spike trains · Monkey cortex · Single neurons

1 Introduction

Even though fluctuations in membrane potential are the fundamental signal of single neurons, the action potentials have been and remain a focus of attention in studies into the neural basis of behavior because the action potentials provide the means to transmit a signal over distances of millimeters and further (even meters from brain to spinal cord in tall or long mammals). The sequence of action potentials is a sampling of the membrane potential, with spikes being closer together when the membrane potential is near the firing threshold and further apart when the membrane is farther from the firing threshold. The membrane potential is driven by summing over the neuron's inputs, which themselves are most generally the consequence of action potentials arriving at the input axonal terminals.

By accepting that the spike train represents long distance transmission of the information encoded by the membrane potential, it is possible to simplify the comparison of behavior to the information encoded in spike trains, so as to overlook a

B.J. Richmond (✉)
Section on Neural Coding and Computation, Laboratory of Neuropsychology, Department of Health and Human Services, US National Institute of Mental Health, Bethesda, MD 20892, USA
e-mail: bjr@ln.nimh.nih.gov

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_3,

detailed consideration of the biophysics. As a disclaimer, it must be understood that this is not meant in any way to minimize the critical importance of the biophysics, but rather represents a simplification that provides a tractable means for correlating single neuronal signals and behavior.

This approach provides a simple information channel model for evaluating neuronal responses. The spike train is the carrier, and the modulations of it represent the information carried by the channel. To evaluate the information it is necessary to learn what the code elements in the channel are and how they are modulated by the channel inputs. Despite a considerable amount of work, the "code" is not certain.

2 Results and Discussion

It is generally agreed that the number of action potentials is a critical feature of the code, i.e., the number of action potentials varies as a function of the conditions. It is straightforward to establish that pattern of the spikes over time also varies as a function of the number of spikes. The number of spikes varies across trials, and the spikes are distributed differently through time. This difference is best illustrated when the shapes of spike densities (a version of the peristimulus time histogram that is well-behaved in the frequency domain) have the same number of spikes but are different in their temporal patterns, with, for example, a burst-pause pattern for one stimulus or condition, and a more steady firing rate for another. Nonetheless, all of the spike trains elicited in one condition are different from the others. This suggests a model in which the spikes are stochastic but the underlying drive is deterministic.

Our estimate of the deterministic pattern of excitation and inhibition arises using the spike density estimated over trials. Through this procedure we make a varying rate estimate or the probability of a spike as a function of time. In this view there is always a stochastic aspect of spike generation. This time-varying estimate improves with larger numbers of trials and spikes.

If we then take the spike train as an ordered set of points sampling the spike density, we end up with a simple statistical model of spike generation. This is the simple model described by order statistics:

$$h_{s,n,k}(t) = \binom{n}{k} kF^{k-1}(t)f(t)[1 - F(t)]^{n-k}$$

where $h_{s,n,k}(t)$ is the unconditional probability of the kth spike out of n spikes at time t for stimulus s, $f(t)$ is the probability of a spike at time t (i.e., the spike density normalized to total 1), and $F(t)$ is the cumulative probability of a spike occurring by time t. $\binom{n}{k}$, that is, $\frac{n!}{k!(n-k)!}$ or n choose k, is needed to make the order statistic a properly normalized probability function summing to 1.

For the order statistic model to be manageable, we assume that rate function over time, that is, the normalized spike density function, is deterministic. It must arise from the interplay between excitatory and inhibitory inputs, and, for any given

condition, we assume that this is constant. So, for example, in the primary visual cortex, V1, it is will be related to the stimulus that is presented. In this model the bandwidth used to construct the spike density function plays a substantial role in the precision with which spike times are determined (after all, we are just sampling this density function). If spikes occur with little time uncertainty from one trial to the next, the spike density will reflect this with a steep rise. The steepness of the rise also clearly depends on the bandwidth of the smoothing function for the spike density. This leads to the conclusion that a smoothing function that adapts its bandwidth locally is desirable to capture reproducible spike events.

In this model the stochastic elements are completely captured in the variability of the spike counts across trials or episodes. This can be measured and we can use our model to generate spike trains with the needed number of spikes (so that we get stochastically generated spike trains). This arises by assuming that the needed number of spikes comes from a binary process with the probabilities given by the spike density function multiplied by the number of spikes needed in the sampling epoch. This model can also be implemented as a spike count matched model. In this model we take advantage of the time-rescaling theorem. We construct the spike density function so that it maps equi-probability steps into time steps. Thus on the probability scale all steps are equal, and generating random numbers from a uniform distribution can be placed in the appropriate bin, that is, at the appropriate time. This relation makes it possible to generate random numbers from a uniform process, a straightforward procedure, and remap them into an arbitrary stochastic rate varying point process.

With a small bit of algebra, the approach can be used to define a simple decoder, that works millisecond-by-millisecond, whether there are spikes at any time point or not. The fundamental equation of order statistics above gives the unconditional probability of the kth spike in a train with n spikes. It is desirable to have the conditional probability of each spike given the history of the spike train at time t, taken over all spike trains no matter how many spikes. The probability of a spike train with n spikes is $p(n) = f(n)/\sum_{n'} f(n')$ where $f(n)$ is the number of times that n spikes occurred in the sample, and $p(n)$ is the probability that a train with n spikes occurs.

Using this, the probability of kth spike for stimulus s across all spike counts, $h_{k,s}(t) = \sum_{n'} p(n) h_{k,n',s}(t)$. Given this, $p(s|t) = \frac{p(s)h_{1,s}(t)}{\sum_{s'} p(s')h_{1,s'}(t)}$ is the probability of stimulus s given the first spike at time t. This can all be done again after the first spike, with the $p(s\,|t\,)$ becoming the new prior probabilities, $p(s)$. For times with no spikes, the probability is $(1 - p(s\,|t\,))$. This, then, is a complete decoder for spikes as they arrive or don't, millisecond-by-millisecond. It straightforward to extend this model to sets of neurons, but their utility is limited by the difficulty in measuring joint spike count probability functions across neurons in experimental data. From this approach we gain the opportunity to compare neural responses to behavioral performance.

Acknowledgments This research was supported by the Intramural Program of the National Institute of Mental, U.S. National Institutes of Health.

Bibliography

1. Optican, L.M., Richmond, B.J.: Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. J. Neurophysiol. **57** (1987) 162–178.
2. Oram, M.W., Wiener, M.C., Lestienne, R., Richmond, B.J.: Stochastic nature of precisely timed spike patterns in visual system neuronal responses. J. Neurophysiol. **81** (1999) 3021–3033.
3. Richmond, B.J., Optican, L.M.: Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. J. Neurophysiol. **57** (1987) 147–161.
4. Richmond, B.J., Oram, M.W., Wiener, M.C.: Response features determining spike times. Neural. Plast. **6** (1999) 133–145.
5. Richmond, B.J., Optican, L.M., Podell, M., Spitzer, H.: Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. J. Neurophysiol. **57** (1987) 132–146.
6. Wiener, M.C., Richmond, B.J.: Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. J. Neurophysiol. **82** (1999) 2861–2875.
7. Wiener, M.C., Richmond, B.J.: Model based decoding of spike trains. Biosystems. **67** (2002) 295–300.
8. Wiener, M.C., Richmond, B.J.: Decoding spike trains instant by instant using order statistics and the mixture-of-Poissons model. J. Neurosci. **23** (2003) 2394–2406.

Phase Transitions in Mesoscopic Brain Dynamics – Implications for Cognition and Consciousness

Hans Liljenström

Abstract Mesoscopic brain dynamics, typically studied with electro- and magnetoencephalography (EEG and MEG) display a rich complexity of oscillatory and chaotic-like states, including many different frequencies, amplitudes and phases. Presumably, these different dynamical states correspond to different mental states and functions, and to study transitions between such states could give us valuable insight in brain-mind relations that should also be of clinical interest. We use computational methods to address these problems, with an objective to find relations between structure, dynamics and function. In particular, we have developed models of paleo- and neocortical structures, in order to study their mesoscopic neurodynamics, as a link between the microscopic neuronal and macroscopic mental events and processes. In this presentation, I will describe different types of models, where the emphasis is on network connectivity and structure, but also including molecular and cellular properties at varying detail, depending on the particular problem and experimental data available. We use these models to study how phase transitions can be induced in the mesoscopic neurodynamics of cortical networks by internal (natural) and external (artificial) factors. We relate and discuss the models and simulation results to macroscopic phenomena, such as arousal, attention, anaesthesia, learning, and mental disorders.

Keywords Cortical network models · Mesoscopic neurodynamcis · Oscillations · Chaos · Noise · EEG · Neuromodulation · Electrical stimulation · Anaesthetics · Cognition · Consciousness

H. Liljenström (✉)
Division of Biometry and Systems Analysis, SLU, SE-75007 Uppsala, Sweden; Agora for Biosystems SE-19322 Sigtuna, Sweden
e-mail: hans.liljenstrom@et.slu.se

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_4,

1 Introduction

Brain structures are characterized by their complexity in terms of organization and dynamics. This complexity appears at many different spatial and temporal scales, which in relative terms can be considered micro, meso, and macro scales. The corresponding dynamics may range from ion channel kinetics, to spike trains of single neurons, to the neurodynamics of cortical networks and areas [1, 2]. The high complexity of neural systems is partly a result of the web of non-linear interrelations between levels and parts, with positive and negative feedback loops. This in turn introduces thresholds, lags and discontinuities in the dynamics, and opens up for unpredictable and non-intuitive behaviours of the system [3].

Typical for complex systems in general, and for the nervous system in particular, is that different phenomena appear at different levels of aggregation and spatial (and temporal) scale. New and unpredictable qualities emerge at every level, qualities that cannot be reduced to the properties of the components at the underlying level. In some cases, there is a hierarchical structure of a simple kind, where higher *macro* levels "control" lower ones (c.f. the so-called enslaving principle of Haken [4]). However, there could also be a more "bottom-up" interpretation of systems, where indeed the *micro* phenomena, through various mechanisms, set the frame for phenomena at higher structural levels. This interplay between micro and macro levels is part of what frames the dynamics of systems. Of special interest is the *meso* level, i.e. the level in between the micro and the macro, as this is where bottom-up meets top-down [3, 5].

The activity of neural systems often seems to depend on non-linear threshold effects, where microscopic fluctuations may cause rapid and large effects at a macroscopic level. There is a dynamical region between order and pure randomness that involves a high degree of complexity and which seems characteristic for neural processes. This dynamics is very unstable and shifts from one state to another within a few hundred milliseconds or less, typical for chaotic (or pseudo-chaotic) systems.

Despite at least a century of study, the functional significance of the neural dynamics at the different levels is still not clear. Very little is also known about the relation between activities at the different levels. However, it is reasonable to assume that different dynamical states correlate with different functional or mental states. This is also guiding our research and will be discussed more in this presentation.

By studying various kinds of transitions in the brain dynamics, we may be able to reveal fundamental properties of the brain and its constituents, also relating to mental processes and transitions. Such transitions could, for example, involve various cognitive levels and conscious states, which could be of interest, not only to neuroscience, but to psychology, psychiatry and medicine.

2 Mesoscopic Brain Dynamics

The main focus here is on phase transitions at the mesoscopic level of neural systems, as it constitutes a well-studied bridge between neural and mental processes

[5]. In our description, *mesoscopic brain dynamics* refers to the neural activity or dynamics at intermediate scales of the nervous system, at levels between neurons and the entire brain. It relates to the dynamics of cortical neural networks, typically on the spatial order of a few millimetres to centimetres, and temporally on the order of milliseconds to seconds. This type of dynamics can be measured by methods, such as ECoG (electrocorticography), EEG (electroencephalography) or MEG (magnetoence-phalography).

We consider processes and structures studied with a microscope or microelectrodes as a *microscopic* scale of the nervous system. It could, for example, refer to ion channels or single neurons. The *macroscopic* scale, in this picture, corresponds to the largest scale possible to measure with regard to brain activity. Typically, this could concern the dynamics of maps and areas, usually measured with PET or fMRI, or other brain imaging techniques.

Mesoscopic brain dynamics with its transitions, is partly a result of thresholds and the summed activity of a large number of elements interconnected with positive and negative feedback. It is also a result of the dynamic balance between opposing processes, influx and efflux of ions, inhibition and excitation etc. Such interplay between opposing processes often results in (transient or continuous) oscillatory and chaotic-like behaviour [2, 6–8].

The mesoscopic neurodynamics is naturally influenced by the activity at other scales. For example, it is often mixed with noise, generated at a microscopic level by spontaneous activity of neurons and ion channels. It is also affected by macroscopic activity, such as slow rhythms generated by cortico-thalamic circuits or neuromodulatory influx from different brain regions. Effects of arousal, attention, or mood, through neuromodulation or other means, could be seen as a top-down interaction from macroscopic activity to mesoscopic neurodynamics.

3 Computational Methods

A computational approach can be seen as a complement to experimental methods in understanding the complexity of neural systems and processes. Computational methods have since long been used in neuroscience, perhaps most successfully for the description of action potentials [9]. When investigating interactions between different neural levels, computational models are essential, and may, in some cases, be the only method we have. In recent years, there is also a growing interest in applying computational methods to problems in clinical neuroscience, with implications for psychology and psychiatry [5, 10–14].

In our research, we use a computational approach to address questions regarding relations between structure, dynamics, and function of neural systems. Here, the focus is on understanding how transitions between different dynamical states can be implemented and interpreted. For this purpose, I present different kinds of computational models, at different scales and levels of detail, depending on the particular issues addressed. The aim is to use a level of description appropriate for the problem addressed.

4 Neural Network Models with Phase Transitions

In this presentation, I will give a few examples of how computational models can be used to study phase transitions in mesoscopic brain dynamics. As examples of internally/naturally induced phase transitions, I present some models with intrinsic noise, neuromodulation, and attention, which in fact, all may be related. In particular, neuromodulation seems to be closely linked to the level of arousal and attention. As examples of externally/artificially induced phase transitions, I discuss electrical stimulation. Both as electric shocks applied directly onto the olfactory bulb and cortex in an experimental setting with animals, and as electroconvulsive therapy applied in a clinical situation in treatment of psychiatric disorders. The final example is a network model testing how certain anaesthetics may act on the brain dynamics through the selective blocking of ion channels.

In all cases, the mesoscopic scale of cortical networks is in focus, with an emphasis on network connectivity. The objective is to investigate how structure is related to dynamics, and how the dynamics at one scale is related to that of another. I will also discuss how structure and dynamics are related to function, with the general notion being that mesoscopic brain dynamics reflects mental states and processes.

Our model systems are paleocortical structures, the olfactory cortex and hippocampus, as well as neocortical structures, exemplified by the visual cortex. These structures display a complex dynamics with prominent oscillations in certain frequency bands, often interrupted by irregular, chaotic-like activity. In many cases, it seems that the collective dynamics of the cortex after external stimulation is a kind of "resonance" between network connectivity (with negative and positive feedback loops), neuronal oscillators and external input.

5 Discussion

Models should always be adapted to the problem they are supposed to address, with an appropriate level of detail at the spatial and temporal scales considered. In general, it could be wise to apply Occam's razor in the modelling process, aiming at a model as simple as possible, and with few (unspecified) parameters. For the brain, due to its great complexity and our still rather fragmented knowledge, it is difficult to find an appropriate level of description and which details to include. Even though the emphasis may be put at different levels, the different models can often be regarded as complementary descriptions, rather than mutually exclusive. At this stage, it is in general not possible to say which models give the best description, for example when trying to link neural and mental processes, in particular with regard to the significance of phase transitions.

While our models are often aimed at mimicking specific cortical structures and network circuitry at a mesoscopic level, in some cases there is less realism in the connectivity than in the microscopic level of single neurons. The reason for this is that the aim in those cases has been to link the neuronal spiking activity with the collective activity of inter-connected neurons, irrespective of the detailed

network structure. Model simulations then need to be compared with spike trains of single neurons, as captured with microelectrodes or patch clamp techniques. In cases, where the network connectivity is in focus, the network nodes may represent large populations of neurons, and their spiking activity is represented by a collective continuous output, more related to LFP or EEG activity.

Even though attempts have been made, it is a struggle to include several levels of descriptions in a single model, relating the activity at the different levels to each other, [5, 10, 12, 14]. In fact, relating different spatial and temporal scales in the nervous system, and linking them to mental processes can be seen as the greatest challenges to modern neuroscience.

In this presentation, I have focused on how to model phase transitions in mesoscopic brain dynamics, relating it to structural and physiological properties, also trying to relate to some functional aspects.

The main question concerns the functional significance of the complex cortical neurodynamics described and simulated above, and in particular, the significance of the phase transitions between various oscillatory states and chaotic or noisy states. I challenge the rather common view that the electrical activity of the brain, as captured with EEG is an epiphenomenon, without any information content or functional significance.

Instead, our computer simulations support the view that the complex dynamics makes the neural information processing more efficient, providing a fast and accurate response to external situations. For example, with an initial chaotic-like state, sensitive to the input signal, the system can rapidly converge to a limit cycle attractor memory state [16, 17, 19]. Perhaps the most direct effect of cortical oscillations could be to enhance weak signals and speed up information processing, but it may also reflect various cognitive functions, including segmentation of sensory input, learning, perception, and attention.

In addition, a "recruitment" of neurons in oscillatory activity can eliminate the negative effects of noise in the input, by cancelling out the fluctuations of individual neurons. However, noise can also have a positive effect, as will be discussed briefly below. Finally, from an energy point of view, oscillations in the neuronal activity should be more efficient than if a static neuronal output (from large populations of neurons) was required.

The intrinsic noise found in all neural systems seems inevitable, but it may also have a functional role, being advantageous to the system. What, then, could be the functional role of the microscopic noise on the meso- and macroscopic dynamics? What, if any, could be the role of spontaneous activity in the brain? A traditional answer is that it generates base line activity necessary for neural survival, and that it perhaps also brings the system closer to threshold for transitions between different neurodynamical states. It has also been suggested that spontaneous activity shapes synaptic plasticity during ontogeny, and it has even been argued that spontaneous activity plays a role for conscious processes [18, 19].

Internal, system generated fluctuations can apparently create state transitions, break down one kind of order to make place for and replacing it with a new kind of order. Externally generated fluctuations can cause increased sensitivity in

certain (receptor) cells through the phenomenon of *stochastic resonance* (SR) [1, 17, 20, 21]. The typical example of this is when a signal with the addition of noise overcomes a threshold, which results in an increased signal to noise relation.

Indeed, simulations with our paleocortical model demonstrate that an increased neuronal noise level can reduce recall time in associative memory tasks, i.e. the time it takes for the system to recognize a distorted input pattern as any of the stored patterns. Consonant with SR theory, we have found optimal noise values for which the recall time reached a minimum [16, 17].

In addition, our simulations also show that neuromodulatory control can be used in regulating the accuracy or rate of the recognition process, depending on current demands. Apparently, the complex dynamics of the brain can be regulated by neuromodulators, and perhaps also by noise. By this flexible control the neural system could be put in an appropriate state for the right response-action dependent on the environmental demand.

The kind of phase transitions discussed here may reflect transitions between different cognitive and mental levels or states, for example corresponding to various stages of sleep, anaesthesia or wake states with different levels of arousal, which in turn may affect the efficiency and rate of information processing.

I believe a combination of computational analysis and modeling methods of the kind discussed here can serve as an essential complement to clinical and experimental methods in furthering our understanding of neural and mental processes. In particular, when concerning the inter-relation between structure, dynamics and function of the brain and its cognitive functions, this method may be the best way to make progress. The study of phase transitions in the brain dynamics seems to be one of the most fruitful approaches in this respect.

References

1. Århem, P., Liljenström, H.: Fluctuations in neural systems: from subcellular to network levels. In Moss, F., Gielen, S., eds.: Handbook of Biological Physics. Amsterdam: Elsevier (2001) 83–129.
2. Århem, P., et al.: Non-linear state transitions in neural systems: from ion channels to networks. In Liljenström, H., Svedin, U., eds.: Micro – Meso – Macro: Addressing Complex Systems Couplings. London: World Scientific (2005) 37–72.
3. Liljenström, H., Svedin, U., eds.: Micro-Meso-Macro: Addressing Complex Systems Couplings. London: World Scientific (2005).
4. Haken, H.: Synergetics: An Introduction. Berlin: Springer (1983).
5. Freeman, W.J.: Neurodynamics: An Exploration in Mesoscopic Brain Dynamics. Berlin: Springer (2000).
6. Freeman, W.J.: Mass Action in the Nervous System. New York, NY: Academic Press (1975).
7. Skarda, C.A., Freeman, W.J.: How brains make chaos in order to make sense of the world. Brain Behav. Sci. **10** (1987) 161–195.
8. Haken, H.: Principles of Brain Functioning. Berlin: Springer (1996).
9. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. **117** (1952) 500–544.
10. Wright, J.J., Liley, D.T.J.: Dynamics of the brain at global and microscopic scales. Neural networks and the EEG. Behav. Brain Sci. **19** (1996) 285–320.

11. Huber, M.T., Braun, H.A., Krieg, J.C.: Consequences of deterministic and random dynamics for the course of affective disorders. Biol. Psychiatry. **46** (1999) 256–262.
12. Robinson, P.A., et al.: Neurophysical modelling of brain dynamics. Neuropsychopharmacology **28** (2003) S74–S79.
13. Steyn-Ross, D.A., et al.: The sleep cycle modelled as a cortical phase transition. J. Biol. Phys. **31** (2005) 547–569.
14. Liljenström, H.: Neural stability and flexibility – a computational approach. Neuropsychopharmacology. **28** (2003) S64–S73.
15. Gu, Y., et al.: Analysis of phase shifts in clinical EEG evoked by ECT. Neurocomputing. **65–66** (2005) 475–483.
16. Liljenström, H., Autonomous learning with complex dynamics. Int. J. Intell. Syst. **10**: (1995) 119–153.
17. Liljenström, H., Wu, X.: Noise-enhanced performance in a cortical associative memory model. Int. J. Neural Syst. **6** (1995) 19–29.
18. Århem, P., Johansson, S.: Spontaneous signalling in small central neurons: mechanisms and roles of spike-amplitude and spike-interval fluctuations. Int. J. Neural Syst. **7**: (1996) 369–376.
19. Århem, P., Liljenström, H., eds.: Beyond cognition – on consciousness transitions. In Consciousness Transitions – Phylogenetic, Ontogenetic and Physiological Aspects. Amsterdam: Elsevier (2007) 1–25.
20. Bulsara, A., et al.: Stochastic resonance in a single neuron model: theory and analog simulation. J. Theor. Biol. **152** (1991) 531–555.
21. Mandell, A.J., Selz, K.A.: Brain stem neural noise and neocortical resonance. J. Stat. Phys. **70** (1993) 355–3373.

Anatomical and Topological Foundations of Cognitive Neurodynamics in Cerebral Cortex

Walter J. Freeman

Abstract Perception requires action of the body into the environment in search of information that the brain needs to optimize engagement of its body with the world. Actions to achieve maximum grip are implemented in sequential motor command transmitted by the limbic system into the brain stem. Simultaneously preafference sends efference copies to the sensory cortices that activate landscapes of chaotic attractors, which serve to predict the sensory consequences of the impending actions. The microscopic sensory impact delivers trains of action potentials to the primary sensory areas. The trains select relevant basins of attraction; selected attractors send wave packets into the limbic system, where they are integrated in time and space into a multisensory percept that is re-transmitted to all sensory areas as efference copies. The action-perception cycle is conceived topologically as a helix for predictive hypothesis-testing and knowledge-accrual. The cycle categorizes sensory inputs, which then choose appropriate actions.

Keywords Action-perception cycle · Chaotic dynamics · Electrocorticogram (ECoG) · Intentionality · Knowledge · Preafference · Phase transition · Topology

1 Introduction

It is instructive to ask why plants have no brains. With rare exceptions plants stay where their seeds fell; animals browse. Their search for food and shelter requires musculoskeletal apparatus, which in turn requires a fast-acting communication network to coordinate the parts. The network is useless without centralized direction and control, which is provided by a brain. The roles of the brain are to sense and assess the current state of its surround that includes its body; to project current trends

W.J. Freeman (✉)
Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720-3206, USA
e-mail: dfreeman@berkeley.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_5,

and assess multiple future probabilities of states of deprivation or satiety; to select that state which is most desirable; to construct a hierarchy of proposed actions constituting a plan to achieve that state; to predict the changes in sensory stimuli that will be consequences of the actions; and to formulate commands to the motor systems and corollary discharges to the sensory systems in sequences as the selected course of actions unfolds. This is what brains do. The question that neurobiologists try to answer is, how do they do it?

Fortunately the same basic mechanisms hold for both simple behaviors, such as a rat pressing a bar for food, and complex behaviors, such as a farmer negotiating a bank loan and a statesman negotiating a treaty for world peace. All goal-oriented behaviors involve perception and action in cycles, so that each perception becomes the basis for the next action, and each action leads to new perception (Fig. 1). Cognitivists commonly break the loop outside the body in order to study stimulus-response relations in the paradigms of classical and operant conditioning, because they afford efficient behavioral control [1] Philosophers open the loop inside the brain by conceiving action as leading to perception, because the primacy of intention captures the spontaneity and inner-directedness of self-organized behavior [2]. Neurobiologists combine both perspectives in their efforts to learn how populations of brain neurons coordinate their activities to produce behaviors that engage brain, body, and environment. These behaviors are intentional in one of three senses: the psychological sense that behavior is goal-directed, or the cognitivist sense that brains connect symbols with the objects they represent (*aboutness* [3]), or the phenomenological sense that an organism brings its life history to each task [4].

These three meanings of the word *intentional* are the product of three differing paradigms [5] that comprise current neuroscience: physiological psychology, which is dominated by stimulus-response determinism; computational neuroscience, which relies on explanations using rule-driven symbol manipulation by neurons in networks; and dissipative brain dynamics [6] which conceives brain activity in terms of chaotic patterns constructed by self-organized populations of neurons [4, 7, 8]. The paradigms differ in classic experiments, rules of proof and evidence, and the meanings of shared terminology, so they cannot disprove one another, but they can and do complement each other. None is wholly true or false; each is useful in one or another area of neuroscience. The question is: which paradigm is applicable to which dynamics?

My aim here is to describe the flows of neural activity through the vertebrate brain in the most general terms, which is the topology of neurodynamics based on functional anatomy. Barrett [9] wrote: " ... the answer to the question of what is the most fundamental physical description is that it is a description of the topology of the situation. With the topology known, the group theory description is justified and equations of motion can then be justified and defined in specific differential equation form. If there is a requirement for an understanding more basic than the topology of the situation, then all that is left is verbal description of visual images" (p. 3). By stepping back from the bench work of collecting and modeling data one can ask, what are the irreducible differences among alternate descriptions of the neurodynamics of cognition? In particular I will contrast

environmental determinism as posited by the reflexive and computational paradigms with the paradigm of self-determinism by chaotic brain dynamics [8] and describe topological differences.

2 Anatomy and Topology Among Sensory and Limbic Cortices

There is no serious doubt that the forebrain is the organ of intentionality. Reflex responses to sensory input persist after severance of the forebrain from the brain stem, and while the decerebrate animal can walk, trot, gallop and run when placed on a treadmill, it cannot go anywhere intentionally and will starve to death unless food is placed in its mouth triggers reflex chewing. The disconnected cerebrum above the cut usually sleeps. Analyses of the brains of the simplest vertebrates show that three cortical areas in each cerebral hemisphere (Fig. 1) suffice for the simplest intentional behaviors: sensory (mainly olfactory), motor (mainly control of locomotion and the jaws), and associational (the hippocampal formation providing spatial and temporal orientation). The global interaction of these three cortices in conjunction with groups of neurons comprising the basal ganglia in the interior of the hemisphere form the limbic lobe, which generates an evolving attentive state that constructs the latent possibilities of future behaviors, from which a decisive selection is made by the dynamics of the interactions of brain, body and environment. Humans and other mammals have more complex brains; however, the newly added parts – the neocortex, neostriatum, and neothalamus – are in crucial respects mere elaborations on the topology of flows within the limbic system.

A topological representation of the forebrain (Fig. 2) is derived by showing that the sensory and motor systems operate in parallel in controlling the body and in sensing the world, so the systems are grouped into a motor module and a sensory module. The multisensory integration in mammals is performed in the entorhinal

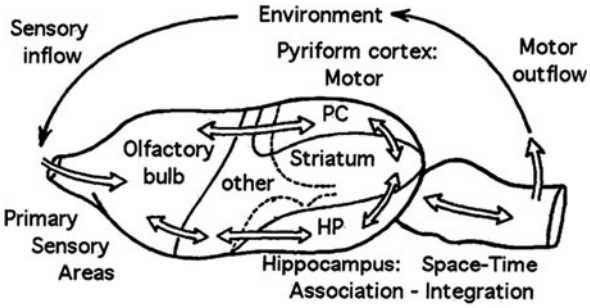


Fig. 1 Each cerebral hemisphere of the simplest brain of vertebrates has a sensory lobe (mainly olfactory), a motor lobe controlling the jaws and body (pyriform), and a hippocampal formation doing spatiotemporal integration of intentional behaviors. From [5]

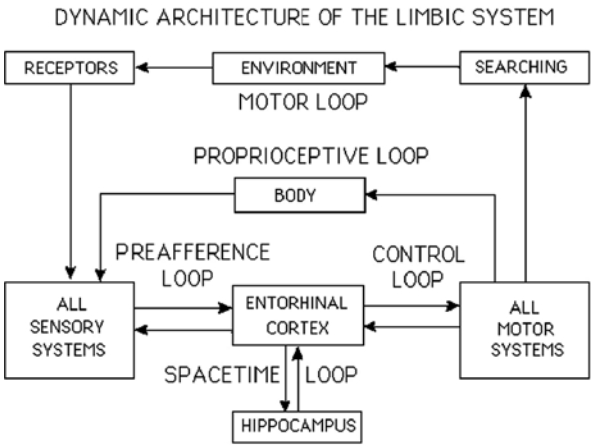


Fig. 2 A schematic shows the topology of the flow of neural activity in the construction of intentional behavior as a collection of *nested circles*. The modalities of sensation and motor output are in parallel and collapsed each in a single node. The two main components of the limbic system form the "space-time loop" that provides orientation in space and sequencing of episodic memories. Closure of the action-perception cycle is through the body and the environment. In engineering models these loops are treated as feedback control loops; physiologically they are homeostatic loops; topologically they are circles. From [5]

cortex, which provides the main input to the hippocampus. It also relays hippocampal output to all sensory and motor areas in preafference, which differs from proprioception that operates through the body and from exteroception that operates through the world. All of these loops can be described topologically as circles. This description is based on studies of the electrocorticogram (ECoG) that document the wide distribution of neural activity in the brain preceding and accompanying intentional actions [4], and on the necessity for integration of brain activity and behavior over time and space. For example, when a hungry organism searches for food, it does so in a sequence of moves. At each move it samples the world by sniffing, listening and looking so as to detect and classify stimuli in all modalities, evaluate their strengths, and hold these values in short-term memory. At the next move, it takes new samples that must be processed and compared with previous sample strengths stored in memory. The organism must also remember where it was at the last sample, what motor command was issued, whether the action commanded actually took place, and where the action was intended to bring the body in respect to past locations and notable landmarks.

These operations are the well-known functions of the hippocampus that provide spatial orientation and temporal sequencing by its cognitive map and short-term memory. The organism needs all this information in integrated form in order to determine what to do next: continue, go left, go right, or reverse. This example is about as simple as one can conceive for a goal-directed action, so it is useful as a basis for diagramming the topology before getting deeper into the technical details of how brains use their bodies to perform searches so rapidly and effectively.

3 The anatomy and Topology of Information Flows in Sensation

Full understanding of the issues that surround the foundations of computational and intentional neurodynamics requires a still deeper level of description of the stage at which sensory input is processed by sensory cortex, leading to perception. Again the simplest model is provided by the olfactory system, in which the receptor axons synapse on the olfactory bulbar projection neurons; the axons of those neurons synapse on the pyramidal cells in the pyriform cortex. Only two synapses intervene between odorant receptors and the pyriform output to the brain stem, but with a major topological difference in connections between the two synaptic layers. The connection from receptors to bulb provides topographic mapping (Fig. 3), like that in all other exteroceptor systems. The connection from bulb to pyriform is through a divergent-convergent projection that enacts a spatiotemporal integral transform. Every bulbar neuron projects broadly to the cortex; every receiving neuron gets input from a broad distribution of bulbar neurons. While input to the bulb is carried by axons conveying amplitudes by pulse frequencies, bulbar output is expressed in distributions of pulse densities. The only activity that does not undergo smoothing and attenuation by this transform is that which has approximately the same instantaneous frequency and phase in the entire bulb. That activity is restricted to what the bulb constructs within itself. Sensory evoked activity lacks this spatial coherence in response to reception of input by the cortex, so it is attenuated by smoothing in the transmission of cortical output.

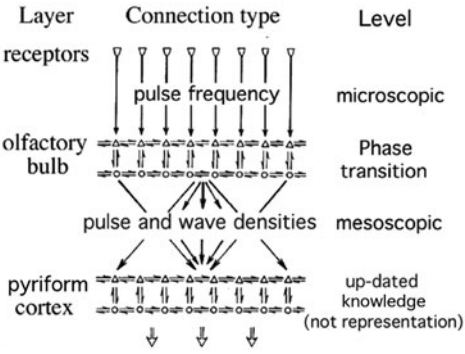


Fig. 3 The geometry is schematized of the transition from sensation to perception. The anatomical projection of the olfactory receptors to the bulb provides the basis for topographic mapping of input onto the bulbar surface. The bulbar projection neurons broadcast their output to widely distributed cortical neurons, which integrate pulses from widely distributed bulbar neurons. This pathway performs a continuous spatiotemporal integral transformation resembling that of a holograph but without an inverse and with irretrievable loss of information through many-to-one convergence and smoothing, which are necessary for abstraction. Neocortical pathways have multiple relays intervening between receptor and cortex, but with preservation of topographic mapping to and to some extend beyond the first cortical synapse, but with divergence-convergence there and thereafter

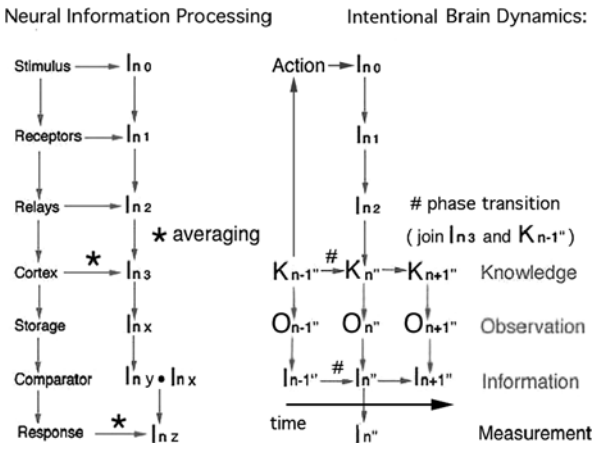


Fig. 4 *Left:* The topology is shown of the transition from sensation to perception. Two models for neural information processing are compared: NIP and IBD. The key operation in NIP is time-locked averaging (*) of sequential patterns. Each sampling is initiated by an input CS; transmission is feedforward. The topological representation is a *line* segment. *Right:* The key operation in IBD is the phase transition (#), through which a selected pattern is created from knowledge Kn-1" and modified by incorporation of the new sensory information In3 that up-dates the knowledge of the stimulus and its meaning [1] to Kn". The action-perception cycle is determined to be an open, feed-forward cycle in time and neither an arc nor a homeostatic feedback loop. The topology is a helix with time along the axis. Convergence to an attractor is many-to-one, irreversible, information-destroying; hence the convergence constitutes a topological obstruction. There is no going back in time

The premise for both stimulus-response and computational models of cognition is that sensory stimuli provide information that is extracted by the specialized sensory receptors and is injected through relays into cortex by trains of action potentials. There it is postulated to be further processed, stored, and retrieved for comparison with fresh information as the basis for stimulus classification and decision-making. The information in sensory cortical output is thought to be delivered to motor cortices, from which it is relayed to the musculoskeletal system for a response, which may or may not invoke reinforcement. The topology of the flow is a line segment (Fig. 4, NIP); in this sense these are linear models.

The evidence comes largely from action potentials that have been recorded from neurons along pathways from receptors to the sensory cortices and within the cortices. The evidence is commonly interpreted as showing that the features of stimuli are *represented* by neural firings of *feature detector* neurons, which are combined by averaging after phase locking of the trains of action potentials *(feature binding*) that represent objects [10] This model requires the assumption of stationarity and fixity of the firing frequency at which the phases are defined, so that responses can be averaged (* in Fig. 4, NIP) over multiple presentations. Averaging reduces the variance caused by the on-going background activity of cortex, which is treated as 'noise'. However, the background activity is not stationary [11]. Furthermore sensory cortices do not have the neural machinery required to accumulate multiple responses

over repeated trials and average them. The ergodic hypothesis, that the output of a population of neurons at any one moment can be represented by the output of any neuron observed sufficiently often, does not apply to feature detector neurons. They are not identical like molecules, nor do they function in autonomous isolation. Thus stimulus-dependent information processing works well to model information from receptors (Ino) to cerebral cortex (In3), but they leave a large gap (* – * in Fig. 4,) between the cortical activity of single cells driven by stimuli (Inx) and the information in conditioned responses to conditioned stimuli (Inz).

4 The Anatomy and Topology of Flows in Perception

That gap can be filled by a model that has been derived using intentional brain dynamics (Fig. 4, IBD) based on recordings of the patterns of dendritic field potentials (the electrocorticogram, ECoG [12, 13]) in sensory cortical responses to conditioned stimuli. These patterns lack the invariance with respect to stimuli that the feature-binding hypothesis requires. Instead the observed patterns change with the history of reinforcement and the context in which the stimuli are received [11, 13–15]; their information content is defined not solely by the information in the stimulus but by context and experience with reinforcements. The memory of the experience is embedded in the rich texture of modified synapses distributed broadly in each sensory cortex. Two synaptic connectivity patterns are inferred from measurement and classification of spatial patterns of amplitude modulation (AM) of carrier waves in the gamma range [7]. An immediate change in pattern is observed during acquisition of a conditioned response (CR) to a reinforced conditioned stimulus (CS). That early change manifests the construction in sensory cortex of a Hebbian nerve cell assembly by strengthening the connections of co-activated neurons on each CS presentation and building a correlation map over multiple trials. The assembly thereafter provides for generalization to the class of the CS over any combination of activated receptors. A delayed change appears one or more days later by consolidation, as revealed by the further evolution of the AM pattern to a new and more stable classifiable form, in which all cortical neurons participate, either by reducing or increasing their firing rates in the construction and maintenance of the AM pattern covering the entire sensory cortex.

The delayed change manifests the construction by synaptic modification of a chaotic attractor in the state space of cortical dynamics [7, 8]. The Hebbian assembly defines the basin of attraction and provides the synaptic connection path to access the attractor. The widely distributed modified synapses embody the knowledge (Kn”, Fig. 4) that the cortex holds in memory about the stimulus. The observation On” of the AM patterns provides access to the neural activity in which the knowledge is expressed. Measurement of the AM patterns gives the numbers by which the information (In”) used for classification is derived. The important point here is that sensory information (Ino to In3) is absorbed by cortex into its knowledge base, so that the information (In”) that emerges from cortex is unique to each subject and cannot be derived logically from the information (Ino) that was introduced

through the sensory receptors. Each understanding (Kn") results from the construction of new information from prior knowledge (Kn – 1) by expenditure of metabolic energy in cortex, so that the quantity of information in In" can greatly exceed that in In despite losses from generalization, abstraction, and smoothing.

Each CS that a subject can discriminate accesses a chaotic attractor and its corresponding AM pattern [8]. When the subject searches for information by an act of observation, a landscape of attractors is pre-activated by the limbic system through preafference (Fig. 2). The landscape embodies the set of predicted outcomes of the act. The CS selects the basin of the appropriate attractor by exciting its Hebbian assembly, which initiates the construction of an AM pattern using the stored knowledge about the CS (Kn – 1") that is up-dated by the incoming information (In3) to the latest knowledge (Kn"). The next input In3 that up-dates Kn" advances the knowledge to Kn + 1", with accompanying changes from On − 1" through On" to On + 1" and from In − 1" through In" to In + 1". This topological representation shows that the action-perception cycle is not a closed feedback loop. It is a predictive, open, feedforward cycle in the form of a helix with advancement along its axis in time.

The linear models suffice to describe the feedforward dynamics of the sensory input from receptors (Ino) to cortex (In3), and also to describe the feedforward transmission of motor output, but they must be complemented by the homeostatic feedback loops shown in Fig. 2. The most complex and crucial step in the model of intentional dynamics is the transition that takes place upon the injection of receptor axon input pulses into the olfactory bulb (Fig. 3). At this point the flow of microscopic activity that is measured by pulse frequencies on axons selected by the CS is organized by the formation of the mesoscopic AM pattern. That construction is initiated by a break, a discontinuity in the ECoG caused by the flow of dendritic current observed in brain waves, which is labeled *phase transition* [11] in Fig. 3 and signified by (#) in Fig. 4. The flow diagram there highlights serial processing in steps, which are sequences of temporally discrete AM patterns in frames. In this respect the information processing can be described as "cinematographic" [16] in contrast to the common assumption in models based on unit recording and the statistics of point processes that brain dynamics is continuous in time. Each frame is introduced by an abrupt change in cortex by which an orderly pattern emerges from disorganized, chaotic activity [8]. The change is comparable to the phase transition in a physical medium, by which a vapor condenses into a liquid with reduction in degrees of freedom upon the emergence of order from chaos [7, 11].

References

1. Skinner, B.F.: Contingencies of Reinforcement; A Theoretical Analysis. New York, NY: Appleton-Century-Crofts (1969).
2. Merleau-Ponty, M.: The Structure of Behavior [Fischer AL, Trans.]. Boston, MA: Beacon Press (1942/1963).
3. Searle, J.R.: Intentionality. Cambridge, UK: Cambridge U.P. (1983).

4. Freeman, W.J.: Societies of Brains: An Examination of the Neurophysiological Basis of Love and Hate. Mahwah, NJ: Lawrence Erlbaum Assoc (1995).
5. Kuhn, T.S.: The Structure of Scientific Revolutions. Chicago, IL: University Chicago Press (1962).
6. Freeman, W.J., and Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Phys. Life Rev. **3** (2006) 93–118. http://repositories.cdlib.org/postprints/1515
7. Freeman, W.J.: How Brains Make up Their Minds. New York, NY: Columbia U.P. (2001).
8. Skarda, C.A., Freeman, W.J.: How brains make chaos in order to make sense of the world. Brain Behav. Sci. **10** (1987) 161–195.
9. Barrett, T.P.: Topology and the physical properties of electromagnetic fields. Apeiron. **7** (2000) 3–11.
10. Singer, W., Gray, C.M.: Visual feature integration and the temporal correlation hypothesis. Annu. Rev. Neurosci. **18** (1995) 555–586.
11. Freeman, W.J., Kozma, R.: Freeman's Mass Action. Encyclopedia for Computational Neuroscience (2009). http://www.scholarpedia.org/article/Freeman's_Mass_Action
12. Freeman, W.J.: Mass Action in the Nervous System. New York, NY: Academic Press (1975/2004). http://sulcus.berkeley.edu/MANSWWW/MANSWWW.html
13. Freeman, W.J.: Origin, Structure, and Role of Background EEG Activity. Part 3. Neural Frame Classification. Clin. Neurophysiol. **116**(5) (2005) 1118–1129. http://repositories.cdlib.org/postprints/2134/
14. Freeman, W.J., Holmes, M.D., West, G.A., Vanhatalo, S.: Dynamics of human neocortex that optimizes its stability and flexibility. J. Intell. Syst. **21** (2006) 1–21. http://repositories.cdlib.org/postprints/2385
15. Ohl, F.W., Scheich, H., Freeman, W.J.: Change in pattern of ongoing cortical activity with auditory category learning. Nature. **412** (2001) 733–736.
16. Sacks, O.: In the river of consciousness. New York Review. 51 Nr 1, January 15, (2004).
17. Barham, J.: A dynamical model of the meaning of information. Biosystems. **38** (1996) 235–241.

Metastability of Mean Field Neuropercolation – The Role of Inhibitory Populations

Robert Kozma and Marko Puljic

Abstract Critical properties of dynamic models of neural populations are studied. Based on the classical work of Erdős-Rényi on the evolution of random graphs and motivated by the properties of the cortical tissue, a new class of random cellular automata called *neuropercolation* has been introduced recently. This work analyzes the role of inhibitory populations in generating multistable dynamics near criticality. The results are interpreted in the context of experimentally observed meta-stable behavior in the cortex.

Keywords Random graphs · Cellular automata · Neuropercolation · Cortex · Multistability

1 Introduction

Experimental studies indicate that intermittent synchronization across large cortical areas provides the window for the emergence of meaningful cognitive activity in animals and humans [1–3]. In neural tissues, populations of neurons send electric currents to each other and produce activation potentials observed in EEG experiments. While single unit activations have large variability and do not seem synchronous, the activations of neural groups often exhibit synchrony. Various studies have been aimed at the development of models that interpret these experimental findings. Successful approaches include ordinary differential equations with distributed parameters and partial differential equations; see, e.g., [4, 5]. Alternative models are based on neuropercolation, which is the objective of this work. Our work is based on the mathematical theory of random graphs, where the process noise and connectivity density are order parameters instantiating state transitions [6, 7]. Earlier works showed the critical role noise plays in the generation of phase transitions in neuropercolation models [8, 9]. Neuropercolation models improve understanding

R. Kozma (✉)
Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, USA
e-mail: rkozma@memphis.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_6,

of the relationship between the system's stochastic components and its emergent collective behavior. In this work we study neuropercolation with populations of excitatory and inhibitory units. Various configurations are used to derive conditions for the onset of phase transitions and multistable oscillations near criticality. The conclusions are beneficial for the interpretation of multichannel EEG data obtained in human cognitive experiments.

2 Overview of Mean-Field Neuropercolation

Standard percolation theory has been an active area of research at the interface of probability theory, combinatorics and physics [10]. Many percolation problems exhibit phase transitions with critical probability p_{crit}. Phase transitions indicate that for p less than p_{crit} only finite clusters exist, and for $p > p_{crit}$ infinite clusters almost surely exist. Here we focus on mean field neuropercolation models following [11]. Let Λ denote the neighborhood of the point of origin $\mathbf{0} = (0, \ldots, 0)$, which is finite subset of $\mathbb{Z}^d$. Let $p\colon 2^{\Lambda} \to [0,1]$ be a function that assigns for each subset $S \subseteq \Lambda$ a probability p_S. Assume p_S depends only on the cardinality of S and whether or not the site itself is active ($z \in S$), which are called isotropic models. We write $p_r^{(0)}$ in place of p_S when $|S| = r$ and $z \notin S$ and $p_r^{(1)}$ when $|S| = r$ and $z \in S$. The model is fully isotropic if $p_r^{(0)} = p_r^{(1)} = p_r$ for all r. In this case, the site itself is treated on the same basis as its neighbors. Finally, we call an isotropic model symmetric if $p_r^{(1)} = 1 - p_{|\Lambda|-r}^{(0)}$ for all r.

In the mean field model, instead of taking $|\Lambda| - 1$ specified neighbors, we take $|\Lambda| - 1$ elements of the grid at random (with replacement). It is clear that the mean field model does not depend on the topology of the grid, and the only information of relevance in X_t is given by its cardinality $|X_t|$. We define x_t to be $|X_t|/N$ where N is the size of the finite grid or torus. Thus $x_t \in [0,1]$ gives the density of points in X_t. If there are N points in the d-dimensional torus then in the mean field model x_t is given as a random process [11]:

$$Nx_{t+1} = |X_{t+1}| = B(Nx_t, f_m^{(1)}(x_t)) + B(N(1-x_t), f_m^{(0)}(x_t)) \tag{1}$$

Here $B(k,p)$ is a binomial random variable giving the sum of k independent Bernoulli random variables, each of which is 1 with probability p. Depending on the actual form of functions $f_m^{0/1}(x)$, the stochastic process $x(t)$ can have various dynamics. It has been shown in [11] that the mean field model gives at least one fixed fixed point, and can have other fixed points, limit cycles and strange attractors. For sufficiently large lattices x_t is normally distributed. The mean value is given by:

$$f_m(x) = x f_m^{(1)}(x) + (1-x) f_m^{(0)}(x) = \sum_r \binom{|\Lambda|}{r} p_r x^r (1-x)^{|\Lambda|-r}, \tag{2}$$

where $p_r = (rp_r^1 + (|\Lambda| - r)p_r^0)/|\Lambda|$. Let us consider the symmetric fully isotropic mean field model on the 2-dimensional lattice. It can be shown that there is a stable fixed point for $p_c < p \leq 0.5$, while there are two stable and an unstable fixed points for $p < p_c$. Here p_c is the critical value of probability p. The fixed point is determined by the condition that $x_{t+1} = f_m(x_t)$. Using the majority update rule in evaluating Eq. (2), we readily arrive at the condition:

$$x = (1-p)\Big(\sum_{r=0}^{\lfloor|\Lambda|/2\rfloor}\binom{|\Lambda|}{r}x^{|\Lambda|-r}(1-x)^r\Big) + p\Big(\sum_{r=\lfloor|\Lambda|/2\rfloor+1}^{|\Lambda|}\binom{|\Lambda|}{r}x^{|\Lambda|-r}(1-x)^r\Big). \quad (3)$$

On the 2-dimensional lattice, the condition for stable solution is readily obtained as $p = 0.5 - 2^{|\Lambda|-2}/(|\Lambda|\binom{|\Lambda|-1}{(|\Lambda|-1)/2})$. After substituting the value $x = 0.5$ at criticality, considering $|\Lambda| = 5$, we get $p_c = \frac{7}{30}$. Near the critical point, the density versus p relationship approximates the following power law behavior with very good accuracy: $|x - 0.5| \propto (p_c - p)^\beta$, where $\beta \approx 0.5$. Figure 1 illustrates the stable density values as solid lines according to Eq. (3). Density level 0.5 is the unique stable fixed point of the process above the critical point $p_c \leq p$, while it becomes unstable below p_c.

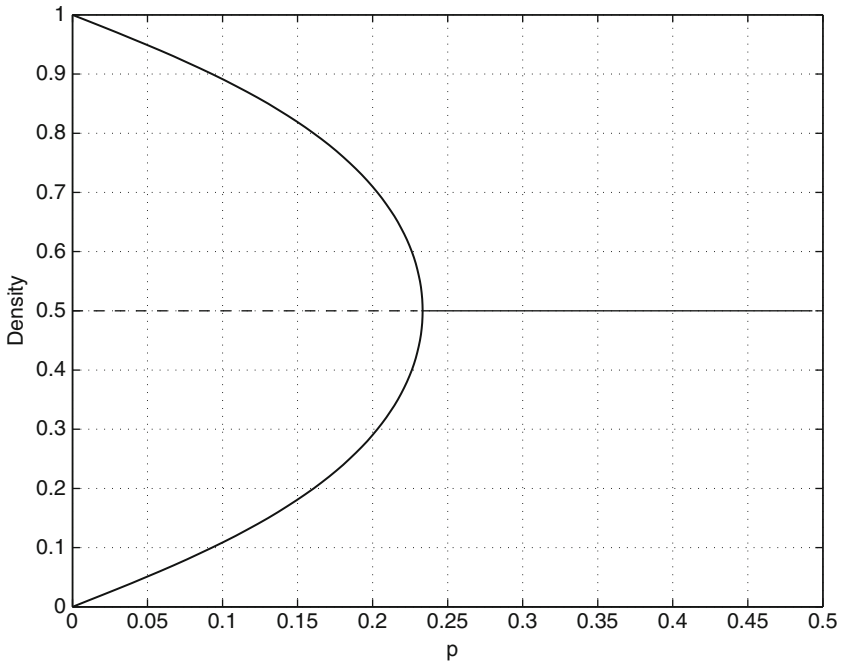


Fig. 1 Activation density x of the mean field model as the function of noise level p. *Solid line*: stable fixed point, *dash* – unstable fixed point; critical probability $p_c = 7/30$

3 Inhibition in Mean Field Approximation

Inhibition is modeled by replacing a connection of an excitatory node by a link from a node from the inhibitory layer. In the case of inhibition, the majority rule is modified as follows. If the inhibitory node is inactive at a certain time instant, its effect gives zero input to the excitatory node into which it has a projection. On the other hand, if the inhibitory node is active, it reduces by one unit the total activation calculated by the summation of the excitatory neighbors of the given node. In the presence of inhibition, the maximum value of sum of neighbors activation is reduced by 1, so the chance that the majority rule produces active site at the next iteration in the excitatory layer is reduced.

First it is assumed that the inhibitory layer is active or inactive independently of the condition of the excitatory layer. This assumption allows to derive analytical expressions for the fixed points. Later this assumption will be relaxed in numerical studies. Using the independence assumption, we modify the condition of fixed point in mean field approximation without inhibition, as given by Eq. (3). The modified equation for the fixed points of the mean field model, in the presence of inhibition writes:

$$\begin{aligned} x = (1-\beta)[&(1-p)(\textstyle\sum_{r=0}^{\lfloor |\Lambda|/2 \rfloor} \binom{|\Lambda|}{r} x^{|\Lambda|-r}(1-x)^r) \\ &+p(\textstyle\sum_{r=\lfloor |\Lambda|/2 \rfloor+1}^{|\Lambda|} \binom{|\Lambda|}{r} x^{|\Lambda|-r}(1-x)^r)] \\ \beta[&(1-p)(\textstyle\sum_{r=0}^{\lfloor |\Lambda|/2-1 \rfloor} \binom{|\Lambda|-1}{r} x^{|\Lambda|-r-1}(1-x)^r) \\ &+p(\textstyle\sum_{r=\lfloor |\Lambda|/2 \rfloor}^{|\Lambda|-1} \binom{|\Lambda|-1}{r} x^{|\Lambda|-r-1}(1-x)^r)]. \end{aligned} \tag{4}$$

The term with factor $1-\beta$ describes the contribution of sites without inhibition, while the term with β is the modified contribution in the presence of inhibition [12]. Eq. (4) gives the fixed point value of x as the function of p and β. Results of numeric analysis are illustrated on Fig. 2. One can see the drastic change of the solutions near criticality as compared to Fig. 1. In the absence of inhibition $\beta = 0$ we observe 2 stable and one unstable fixed points, while the behavior is more complex when $\beta > 0$ and the stable and unstable fixed points are intertwined. Evaluations with more detailed inhibitory models show the occurrence of additional fixed points leading to multistability and to the onset of oscillatory behaviors.

4 Discussion and Conclusions

Random cellular automata have been proven to be very useful tools for modeling cortical neurodynamics. Using the concept of Freeman K models, the implementation of K0 and KI sets has been successfully completed in recent years using neuropercolation models [8, 11]. In a single layer KI set, mutual excitation of

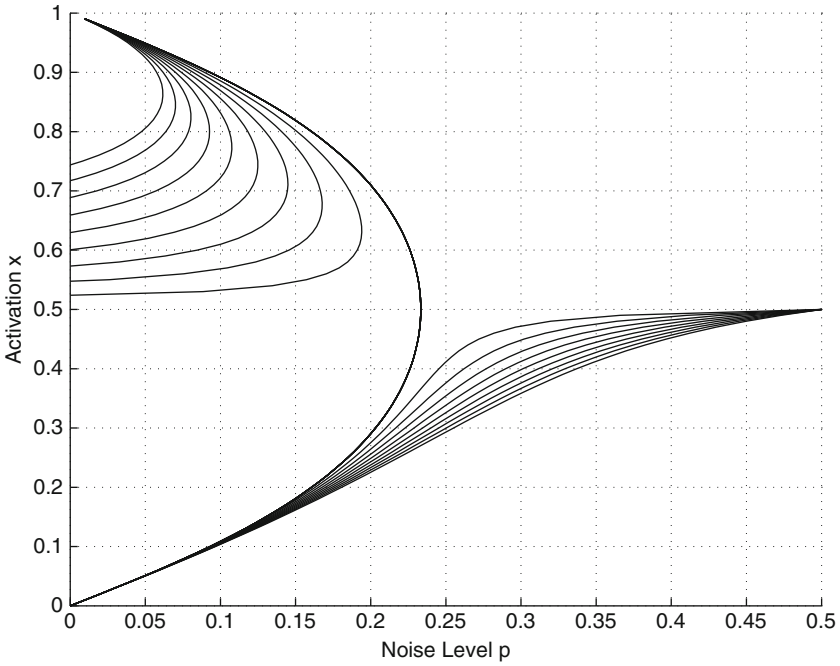


Fig. 2 Activation density x of the mean field model as the function of noise level p. Different curves correspond to increasing beta from 0 to 1 with increments 0.1

excitatory nodes provides the sustained broad-band activity. The activity of a single excitatory layer is self-stabilized by a non-zero point attractor, giving rise to very broad band noise. Neuropercolation models display complex behaviors which are difficult to analyze rigorously. Computer simulations of large-scale lattices have been used successfully to study lattice dynamics in this case.

As a next step, the effect of inhibitory populations has been analyzed using computer simulation of KII sets [12]. The results demonstrate the presence of two-step transition to criticality, which has been thoroughly analyzed using finite size scaling theory. In the coupled layers with appropriate topology and noise level, the inhibitory layer contributes negative feedback, leading to narrow-band oscillations.

The present work gives a theoretical approach to study narrow-band oscillations in neuropercolation models of KII sets. It is shown that the system can exhibit multistable behavior leading to the onset of narrow-band oscillations. Quantitative evaluation of the critical probability and other dynamical properties of neural populations indicate the advantages of the introduced neuropercolation model in the interpretation of phase transitions in cortical tissues.

Acknowledgments This research has been funded in part by AFOSR (Dr Jun Zhang) program on mathematical modeling of cognition and decision, and FedEx Institute, Memphis, Technology Research Grant.

References

1. Freeman, W.J.: Tutorial on neurobiology – From single neurons to brain chaos. Int. J. Bifurcat. Chaos. **2**(3) (1992) 451–482.
2. Freeman, W.J.: How Brains Make Up Their Minds. London: Weidenfeld and Nicolson (1999).
3. Stam, C.J.: Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. Clin. Neurophysiol. **116** (2005) 2266–2301.
4. Kozma, R., Freeman, W.J.: Chaotic resonance – Methods and applications for robust classification of noisy and variable patterns. Int. J. Bifurcat. Chaos. **11**(6) (2001) 2307–2322.
5. Steyn-Ross, D.A., Steyn-Ross, M.L., Sleigh, J.W., Wilson, M.T., Gillies, I.P., Wright, J.J.: The sleep cycle modelled as a cortical phase transition. J. Biol. Phys. **31** (2005) 547–569.
6. Erdős, P., Renyi A.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5** (1960) 17–61.
7. Bollobás, B.: Random Graphs. Orlando: Academic Press (1985).
8. Kozma, R., Puljic, M., Balister, P., Bollobás, B., Freeman, W. J.: Phase transitions in the neuropercolation model of neural populations with mixed local and non-local interactions. Biol. Cybern. **92**(6) (2005) 367–379.
9. Puljic, M., Kozma, R.: Activation clustering in neural and social networks. Complexity **10**(4) (2005).42–50.
10. Grimmett, G.: Percolation in Fundamental Principles of Mathematical Sciences. Berlin: Spinger-Verlag (1999) xiv +444pp.
11. Balister, P., B. Bollobás, R. Kozma.: Mean field models of probabilistic cellular automata. Random. Struct. Algorithms. **29** (2005) 399–415.
12. Puljic, M., Kozma, R.: (2008). Narrow-band oscillations in probabilistic cellular automata. Phys. Rev. E. **78** (2008) 026214.
13. Binder, K.: Z. Phys. B. **43** (1981) 119.

Spontaneous Low-Frequency Fluctuation Observed with Functional Magnetic Resonance Imaging as a Potential Biomarker in Neuropsychiatric Disorders

Yuan Zhou, Kun Wang, Yong Liu, Ming Song, Sonya W. Song, and Tianzi Jiang

Abstract As functional magnetic resonance imaging (fMRI) studies have yielded increasing amounts of information about the brain's spontaneous activity, they have revealed fMRI's potential to locate changes in brain hemodynamics in neuropsychiatric disorders. In this paper, we review studies that support the notion that changes in spontaneous low-frequency fluctuation (SLFF) observed by fMRI can be used as potential biomarkers for diagnosis and treatment evaluation in neuropsychiatric disorders. In this paper, we review the methods used to study SLFF from individual region of interest analysis, to local network analysis, to whole brain network analysis. We also summarize the major findings associated with major neurological and psychiatric disorders obtained using these methods.

Keywords Resting-state fMRI · Low frequency fluctuation · Functional connectivity · ICA · Alzheimer's disease · Schizophrenia

1 Introduction

Accurate diagnosis of neuropsychiatric disorders presents a major challenge for psychiatrists and clinicians. The current situation, in which neuropsychiatric disorders are diagnosed mainly based on clinical symptoms and medical history, has spurred the search for objective biomarkers, such as neuroimaging markers.

Spontaneous brain activity, as observed by functional magnetic resonance imaging (fMRI), is termed spontaneous low-frequency fluctuation (SLFF) in blood oxygen level dependence (BOLD) signal and has recently attracted the attention of researchers as a potential biomarker for locating changes in brain hemodynamics associated with diseases because of its clinical advantages, such as non-invasive,

T. Jiang (✉)
National Laboratory of Pattern Recognition, LIAMA Center for Computational Medicine, Institute of Automation, Chinese Academy of Sciences, Beijing, PR. China
e-mail: jiangtz@nlpr.ia.ac.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_7,

easily performed [1, 2]. Findings of the physiological origin of spontaneous brain activity further strengthen the possibility that detected changes in spontaneous brain activity can be used as potential biomarkers (for reviews, see [1, 3]). This has inspired more and more researchers to explore whether endogenous fMRI markers can characterize the neurophysiological changes associated with disease or drug treatments. To date many studies have been performed involving nearly all of the major neurological and psychiatric disorders. Various methods for detecting spontaneous activity have been used to investigate biomarkers that can be related to specific neuropsychiatric disorders. In this paper, we review the methods used to study SLFF from individual region of interest (ROI) analysis, to local network analysis, to whole brain network analysis. We take two representative neuropsychiatric disorders, i.e., Alzheimer's disease (AD) and schizophrenia, as examples to instantiate how the findings obtained using these methods would be helpful to find the biomarkers as diagnosis, therapy evaluation and prognosis of major neurological and psychiatric disorders.

2 Methods and Progress in Understanding Neuropsychiatric Disorders

Correlation or coherence of brain region activity is often thought to reflect functional integration, one of the fundamental principles of functional organization of the brain [4, 5]. Resting-state functional connectivity (rsFC) analysis and independent component analysis (ICA) are two main methods of analysis that can be used to investigate the spatial pattern of spontaneous activity during rest.

Conventional FC measures correlations between a reference time series and another time series [6]. Using this method, fMRI studies have demonstrated that spontaneous BOLD fluctuations are coherent within specific neuro-anatomical systems, such as primary motor auditory, visual cortices, language and limbic systems in healthy subjects (for a review, please see [1]). The rsFC analysis can be divided into three categories: rsFC based on individual regions of interest (ROI), rsFC based on multiple ROIs, and rsFC based on whole brain regions.

The ICA is another commonly used method for identifying spatial patterns in resting-state fMRI data [7–11], This method, like the rsFC based on multiple ROIs, can provide specific information on local brain network. Unlike traditional FC analysis, ICA is a model-free method that decomposes the data into statistically independent components. Such decompositions are very useful because they allow for separation into different coherent resting networks and separate these networks from other effects such as head motion or other physiological confounds (such as cardiac pulsation or the respiratory cycle). The conventional ICA method can only be used for a single subject. Recently, a method for investigating coherent signals at a group level, tensor probabilistic ICA (tensor-PICA), has been proposed [12–14]. This method simultaneously decomposes group fMRI data into modes describing variations across space, time, and subjects. It has been demonstrated that the

Table 1 Methods for investigating spatial patterns of SLFFs and progress in understanding neuropsychiatric disorders

Authors	Disorders	Patient/control	Methods
Greicius et al. [35]	AD, mild	14/14	ICA
Wang et al. [21]	AD, early	13/13	Correlation analysis of individual ROIs
Wang et al. [19]	AD, early	14/14	Correlation analysis of the whole brain and of individual ROIs
Allen et al. [20]	AD, probable	8/8	Correlation analysis of individual ROIs
Sorg et al. [36]	aMCI	24/16	ICA, correlation analysis of individual ROIs, structure
Liang et al. [42]	Schizophrenia	15/15	Correlation analysis of the whole brain
Bluhm et al. [26]	Schizophrenia, chronic and medicated	17/17	Correlation analysis of individual ROIs
Garrity et al. [37]	Schizophrenia	21/22	ICA
Zhou et al. [25]	Schizophrenia, first-episode	17/17	Correlation analysis of individual ROIs
Zhou et al. [39]	Schizophrenia, paranoid	18/18	Correlation analysis of multiple ROIs
Calhoun et al. [38]	Schizophrenia, chronic	20/20	ICA
Zhou et al. [27]	Schizophrenia, paranoid	17/14	Correlation analysis of individual ROIs

tensor-PICA approach can provide useful representations of group fMRI data in resting-state studies [15, 16].

Although the two methods each have their pros and cons [17], both of them are widely used in resting-state fMRI studies of neuropsychiatric disorders. Additionally, some findings using the two methods have yielded consistent results. But some inconsistent results have been reported. In the following section, we will comprehensively review the progress obtained by using each of the two methods from individual ROI analysis, to local network analysis, to whole brain network analysis (Table 1).

2.1 Individual ROI Analysis and Progress in Understanding Neuropsychiatric Disorders

RsFC based on a single ROI is the most common method for investigating resting-state functionality in neuropsychiatric disorders. The key point of this method is to select an appropriate region as ROI. An ideal ROI should satisfy the following criteria: (1) implicated in a pathological lesion or cognitive dysfunction associated with the brain disorder to be investigated; (2) demonstrated as a local abnormality

in anatomy or function in a previous study. Using rsFCs based on an individual ROI has led to some interesting findings for AD and schizophrenia as reviewed below.

AD: Two regions in the default mode network (DMN), the hippocampus and the posterior cingulate cortex (PCC), have attracted the most attention due to their roles in memory and their morphologic and metabolic abnormalities in AD patients [18]. As a core region in the DMN (this network is also called the task-negative network, TNN), the PCC showed decreased correlations with other regions in the same network and decreased anti-correlations with regions in the task-positive network (TPN), which is anti-correlated with the TNN, in early AD compared to healthy controls [19]. Two other recent studies [20, 21] independently investigated the hippocampal rsFC pattern in AD. Both of the studies found that the hippocampus exhibited disrupted rsFCs with extensive regions, including the PCC/PCu. However, Wang and coworkers [21] also found increased rsFCs between the left hippocampus and the right lateral prefrontal cortex, which was absent in Allen's study [20]. The differences in results between the two studies likely arose from differences in the ROIs that were selected, from differences in disease severity [20] and from differences between the sample subjects (for a review, see [22]. In a recent study, healthy subjects exhibited differences in rsFC patterns between the anterior hippocampus and the posterior portion, including the body of the hippocampus and the posterior parahippocampus [23]. Subtle differences in anterior–posterior hippocampal rsFCs may account for the differences in the two studies, which selected different subregions of the hippocampus as ROIs.

Schizophrenia: The dorsolateral prefrontal cortex (DLPFC), because of its local abnormalities in anatomy and function [24] and its role in various neural circuits relevant to the anatomical and physiological mechanisms of cognitive dysfunction in schizophrenia, has attracted the attention of Zhou et al. [25]. By examining the FC pattern of DLPFC in patients with first-episode schizophrenia and matched controls, these researchers from our laboratory found that the DLPFC exhibited decreased rsFC with the PCC, among other regions [25]. Because the regions and functions of the default mode network have been linked with schizophrenia, the PCC, as a core region in this network, was selected as the ROI to compare the SLFF pattern of chronic, medicated schizophrenic patients with that of control subjects [26]. The PCC was found to show decreased rsFCs in regions associated with the DMN including the PCC, lateral parietal, medial prefrontal cortex (MPFC) and cerebellar regions in chronic, medicated schizophrenic patients [26]. Considering that the hippocampus has been implicated as participating in the pathophysiology of schizophrenia, the FC pattern of the anterior hippocampus was also investigated. In patients with schizophrenia, the bilateral hippocampi showed reduced rsFCs to some regions which have been reported to be involved in episodic memory, such as the PCC, the extrastriate cortex, the MPFC, and the parahippocampal gyrus, which may reflect the disconnectivity within a neural network related to the anterior hippocampus in schizophrenia [27].

Comments on this method: rsFC based on an individual ROI can easily be performed and can provide distinct information on specific regions implicated in neuropsychiatric disorders. However, limitations of this method need to be kept in

mind. The method strongly depends on prior knowledge about the related disorders. In published studies, the ROI was selected according to the researchers' interest and outlined by hand [20, 21, 27] selected by software [19, 25] or identified by known coordinates [26, 28] This subjectivity in ROI selection has lead to slight differences in the anatomical location of the ROIs, which led to differences in the ROI FC patterns, as we noted in the case of the discrepancy in the hippocampal FC in AD [20, 21]. In addition, information outside the functional connectivity pattern of a specific ROI cannot be obtained by this method.

2.2 *Local Network Analyses and Progress in Understanding Neuropsychiatric Disorders*

Brain circuits or networks have often been implicated in major neuropsychiatric disorders, such as AD and schizophrenia [29–31], thus the investigation of rsFC within a specific network may improve our understanding of the neural basis of these disorders. This is supported by findings that have revealed the presence of a number of resting-state brain networks, such as the DMN (e.g. TNN) [8, 32, 33], the TPN [32, 33], and the dorsal and ventral attention networks [34] in healthy subjects. Additionally, these networks are known to be related to emotion, memory, attention, and other high brain functions, which have often been observed to be impaired in neuropsychiatric disorders.

AD: Using ICA to isolate the DMN during a simple sensory-motor task, Greicius et al. [35] found decreased connectivity within the DMN in an AD group [35], which is consistent with findings from rsFC based on individual ROIs [19–21]. By applying ICA and ROI-based FC analysis, Sorg et al. further found that the patients with amnestic mild cognitive impairment (aMCI), a syndrome that carries a high risk for developing AD, demonstrated reduced network-related activity in selected areas of the DMN (left PCC and right MPFC) [36]. These studies consistently show that disconnection in the DMN is a distinctive characteristic of AD (Fig. 1). In addition to the DMN, Sorg et al. found the executive attention network to be affected, but the remaining resting networks to be intact, in individuals at high risk for AD. The reduced network-related activity in the executive attention network is in line with observed attentional deficits in MCI and AD, indicating impaired interaction between the two anti-correlated networks (TPN and TNN) that prominently organize intrinsic brain activity [36].

Schizophrenia: Aberrant FC within the DMN has been observed in schizophrenia using ICA [37], demonstrating increased activity in the ACC/MPFC, the parahippocampus and the PCC in patients. These researchers also found that patients with schizophrenia showed significantly more high-frequency fluctuations and controls showed significantly more low-frequency fluctuations in the DMN [37]. This pattern was validated and extended to the rest of the resting-state networks in a later study [38]. By directly investigating the interregional rsFCs among the regions constituting the DMN, increased rsFCs within this network were also observed in

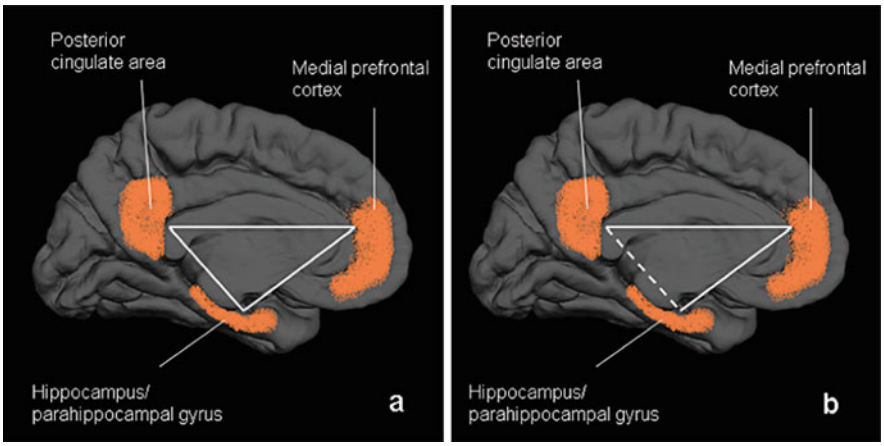


Fig. 1 Altered functional connectivities in the default mode network in Alzheimer's disease. **a** The core brain regions in the default mode network in healthy subjects are illustrated schematically. Prominent components of this network include medial prefrontal regions, posterior regions in the medial and lateral parietal cortex, the lateral temporal cortex and the medial temporal lobe (including the hippocampus and parahippocampal gyrus). Regions within this core brain system are functionally correlated with each other and, prominently, with the hippocampal formation. The *solid line* represents the correlations between the core regions. **b** The functional connectivity between the hippocampal formation and the medial posterior regions were consistently found to be decreased or absent in patients with Alzheimer's disease (for references, please see the main text). The *dashed line* represents decreased or absent correlations between the hippocampus/parahippocampal gyrus and the posterior cingulate area

schizophrenia [39]. The FCs, that primarily increased within the TPN, as well as increased anti-correlations between the two networks were also found in this disease [39]. These findings suggest that these abnormalities related to the DMN could be a possible source for the abnormalities in information processing activities that are characteristic of patients with schizophrenia.

In addition to the TPN and TNN, the frontostriatal system has also attracted the attention of researchers. Salvador et al. used the mutual information measure to compare the resting-state connections among three main components of the frontostriatal system (DLPFC, the basal ganglia, and the thalami) in schizophrenia and in healthy controls. Increased connectivity between the DLPFC and the basal ganglia in schizophrenia were consistently found across low, medium and high frequency bands [40]. These increased connectivities could potentially be responsible for dysfunctions in the frontostriatal loop in patients with schizophrenia.

Comments on this method: Identification of the affected regions is the primary prerequisite for investigations into the rsFC within a disease-related network. However, determining which regions should be recruited for this purpose is limited by the prior knowledge of the researchers. Although some methods, such as ICA, can greatly reduce the researcher's subjectivity, some key steps, such as deciding on the number of components and determining how to classify each component

into noise or physiologically meaningful signals, still depends on the subjective opinions of researchers. This subjectivity leads to differences in the regions that are determined to constitute a network and thus leads to difficulty in comparing results across studies. Furthermore, the functionality and physiological meaning of these resting-state networks need to be further clarified. With this in mind, caution must be exercised in interpreting the findings obtained by comparing the rsFC in the resting-state networks of patients with that of controls.

2.3 Whole Brain Network Analyses and Progress in Understanding Neuropsychiatric Disorders

In contrast to local network-based interregional FC, which focuses on the FCs associated with a few preselected seed regions within a specific network or circuit while ignoring other potentially interesting patterns of connectivity, whole brain network-based FC analysis can objectively and comprehensively detect altered FCs throughout an entire brain level by automatically dividing the entire brain into multiple regions and performing correlation or partial correlation analysis on each pair of these regions. This method was first used in a single patient who was minimally conscious following a brainstem lesion [41] and then was developed and used in AD and schizophrenia [19, 42] (Fig. 2).

AD: Wang et al. [19] found that AD patients show many decreased rsFCs, which are mainly between the prefrontal and parietal lobes, but these patients also show increased rsFCs mainly between regions within lobes, such as within the prefrontal lobe, within the parietal lobe, or within the occipital lobe [19]. These findings

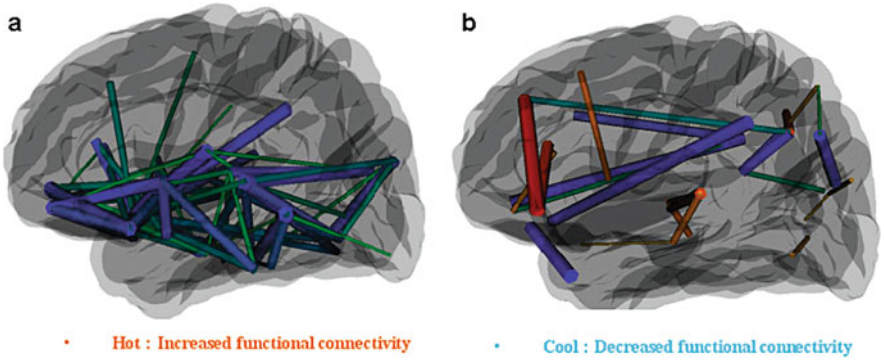


Fig. 2 Altered resting-state functional connectivity in schizophrenia and Alzheimer's disease. **a** Schizophrenia patients mainly showed decreased functional connectivities and such abnormalities were widely distributed throughout the entire brain rather than restricted to a few specific brain regions. **b** Alzheimer's disease mainly showed decreased functional connectivities between the prefrontal and parietal lobes, but increased functional connectivities within the prefrontal lobe, parietal lobe and occipital lobe

are compatible with the anterior–posterior disconnection phenomenon and compensatory effect within lobes observed previously in AD patients (for review, see [22]. More interestingly, by using whole brain network analysis, the authors also found decreased anti-correlations between the two abovementioned intrinsically anti-correlated networks (TPN and TNN), again suggesting that disturbance of the balance between the intrinsically anti-correlated networks may be associated with attention deficits in AD patients.

Schizophrenia: Liang et al. [42] found that in patients with schizophrenia the decreased FCs were widely distributed throughout the entire brain, although most of them were related to the insula, the temporal lobe (including the medial temporal structures), the prefrontal lobe and the corpus striatum. Increased FCs were mainly related to the cerebellum in patients. Although in this preliminary study, interregional anti-correlation was not considered, which made it difficult to directly compare the distribution of these altered FCs with that in other disorders, this study provides further support for the hypothesis that schizophrenia may arise from the disrupted functional integration of widespread brain areas [42].

Comments on this method: By comparing the global distribution of these altered FCs in different disorders, it is possible to find various disease-related characteristics and thus to differentiate different disorders. However, some issues need to be addressed. First, the findings obtained by this method are affected by anatomical parcellation. Therefore, it is necessary to pay attention to the fact that mapping the resulting whole brain network using different templates may induce different findings [43]. Secondly, current automatic registration techniques make it difficult to guarantee the exact match of some small gyri/sulci, especially in the cerebellar lobes, across subjects. Finally, inter-subject variability in the anatomical regions must also be considered carefully in future studies.

3 Conclusions

By analyzing the spatial pattern of spontaneous BOLD activity, some disease-related abnormalities can be obtained. Disrupted rsFCs within the DMN, especially those associated with the hippocampus and the PCC, may be a distinctive characteristic of AD (Fig. 1). The regions showing disrupted rsFCs are highly similar to those that show pathology in the early stages of the disease, as measured by molecular imaging of amyloid plaques using PET, and those that are affected by structural atrophy, as measured by longitudinal MRI [18]. Decreased rsFCs within the attention related-networks is another consistent finding in AD and is consistent with observed attention deficits in MCI and AD. In schizophrenia, aberrant rsFCs within the DMN are found, but the main difference is that the strength of the rsFCs are abnormally increased in patients. By analyzing the clinical correlates of the strength of the rsFC, the rsFCs associated with the component regions of the DMN are found to vary with positive symptoms measured by different clinical scales [26, 27, 37]. The positive symptom-dependent correlation of the component region in the DMN suggests that

functional dysconnectivity in the DMN may be a reflection of an impaired self-monitoring function in schizophrenia, which could lead to positive symptoms such as hallucinations and delusions [44].

In conclusion, altered activity patterns of SLFFs have been found in the major neuropsychiatric disorders, whether at a local level or a global level. These studies highlight the usefulness of resting-state fMRI for studying the brain in neuropsychiatric disorders. More importantly, these studies suggest the possibility that the altered SLFFs in neuropsychiatric disorders could be valuable brain imaging biomarkers for diagnosis, therapy evaluation and prognosis.

Acknowledgments The authors thank Drs. Rhoda E. and Edmund F. Perozzi for English language assistance. This work was supported by the Natural Science Foundation of China (Grant No. 30425004 and Grant No. 30730035), and the National Key Basic Research and Development Program (973) (Grant No. 2007CB512304).

References

1. Fox, M.D., Raichle, M.E.: Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. **8** (2007) 700–711.
2. Bandettini, P.A., Bullmore, E.: Endogenous oscillations and networks in functional magnetic resonance imaging. Hum. Brain Mapp. **29** (2008) 737–739.
3. Kiviniemi, V.: Endogenous brain fluctuations and diagnostic imaging. Hum. Brain Mapp. **29** (2008) 810–817.
4. Zeki, S., Shipp, S.: The functional logic of cortical connections. Nature, **335** (1988) 311–317.
5. Friston, K.J.: Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. **2** (1994) 56–78.
6. Friston, K.J., Frith, C.D., Liddle, P.F., Frackowiak, R.S.: Functional connectivity: the principal-component analysis of large (PET) data sets. J. Cereb. Blood Flow. Metab. **13** (1993) 5–14.
7. Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J.: A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp, **14** (2001) 140–151.
8. Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V.: Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. U S A. **100** (2003) 253–258.
9. Bartels, A., Zeki, S.: The chronoarchitecture of the cerebral cortex. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **360** (2005) 733–750.
10. van de Ven, V.G., Formisano, E., Prvulovic, D., Roeder, C.H., Linden, D.E.: Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum Brain Mapp, **22** (2004) 165–178.
11. Kiviniemi, V., Kantola, J.H., Jauhiainen, J., Hyvarinen, A., Tervonen, O.: Independent component analysis of nondeterministic fMRI signal sources. Neuroimage. **19** (2003) 253–260.
12. Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M.: Investigations into resting-state connectivity using independent component analysis. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **360** (2005) 1001–1013.
13. Beckmann, C.F., Smith, S.M.: Tensorial extensions of independent component analysis for multisubject FMRI analysis. Neuroimage. **25** (2005) 294–311.
14. De Luca, M., Beckmann, C.F., De Stefano, N., Matthews, P.M., Smith, S.M.: fMRI resting state networks define distinct modes of long-distance interactions in the human brain. Neuroimage. **29** (2006) 1359–1367.

15. Damoiseaux, J.S., Rombouts, S.A., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F.: Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. U S A. **103** (2006) 13848–13853.
16. Fransson, P., Skiold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., Aden, U.: Resting-state networks in the infant brain. Proc. Natl. Acad. Sci. U S A. **104** (2007) 15531–15536.
17. Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E.: Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. Neuron. **56** (2007) 171–184.
18. Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L.: The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. **1124** (2008) 1–38.
19. Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., Jiang, T.: Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum. Brain Mapp. **28** (2007) 967–978.
20. Allen, G., Barnard, H., McColl, R., Hester, A.L., Fields, J.A., Weiner, M.F., Ringe, W.K., Lipton, A.M., Brooker, M., McDonald, E., Rubin, C.D., Cullum, C.M.: Reduced hippocampal functional connectivity in Alzheimer disease. Arch. Neurol. **64** (2007) 1482–1487.
21. Wang, L., Zang, Y., He, Y., Liang, M., Zhang, X., Tian, L., Wu, T., Jiang, T., Li, K.: Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI. Neuroimage. **31** (2006) 496–504.
22. Liu, Y., Wang, K., Yu, C., He, Y., Zhou, Y., Liang, M., Wang, L., Jiang, T.: Regional homogeneity, functional connectivity and imaging markers of Alzheimer's disease: A review of resting-state fMRI studies. Neuropsychologia **46** (2008) 1648–1656.
23. Kahn, I., Andrews-Hanna, J.R., Vincent, J.L., Snyder, A.Z., Buckner, R.L.: Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. J. Neurophysiol. **100** (2008) 129–139.
24. Bunney, W.E., Bunney, B.G.: Evidence for a compromised dorsolateral prefrontal cortical parallel circuit in schizophrenia. Brain Res. Brain Res. Rev. **31** (2000) 138–146.
25. Zhou, Y., Liang, M., Jiang, T., Tian, L., Liu, Y., Liu, Z., Liu, H., Kuang, F.: Functional dysconnectivity of the dorsolateral prefrontal cortex in first-episode schizophrenia using resting-state fMRI. Neurosci. Lett. **417** (2007) 297–302.
26. Bluhm, R.L., Miller, J., Lanius, R.A., Osuch, E.A., Boksman, K., Neufeld, R.W., Theberge, J., Schaefer, B., Williamson, P.: Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: anomalies in the default network. Schizophr. Bull. **33** (2007) 1004–1012.
27. Zhou, Y., Shu, N., Liu, Y., Song, M., Hao, Y., Liu, H., Yu, C., Liu, Z., Jiang, T.: Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. Schizophr. Res. **100** (2008) 120–132.
28. Castellanos, F.X., Margulies, D.S., Kelly, C., Uddin, L.Q., Ghaffari, M., Kirsch, A., Shaw, D., Shehzad, Z., Di Martino, A., Biswal, B., Sonuga-Barke, E.J., Rotrosen, J., Adler, L.A., Milham, M.P.: Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. Biol. Psychiatry. **63** (2008) 332–337.
29. Delbeuck, X., Van der Linden, M., Collette, F.: Alzheimer's disease as a disconnection syndrome?. Neuropsychol. Rev. **13** (2003) 79–92.
30. Friston, K.J., Frith, C.D.: Schizophrenia: a disconnection syndrome?. Clin. Neurosci. **3** (1995) 89–97.
31. Friston, K.J.: The disconnection hypothesis. Schizophr. Res. **30** (1998) 115–125.
32. Fransson, P.: Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. Hum. Brain Mapp. **26** (2005) 15–29.
33. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. U S A. **102** (2005) 9673–9678.

34. Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., Raichle, M.E.: Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proc. Natl. Acad. Sci. U S A. **103** (2006) 10046–10051.
35. Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V.: Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc. Natl. Acad. Sci. U S A. **101** (2004) 4637–4642.
36. Sorg, C., Riedl, V., Muhlau, M., Calhoun, V.D., Eichele, T., Laer, L., Drzezga, A., Forstl, H., Kurz, A., Zimmer, C., Wohlschlager, A.M.: Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. Proc. Natl. Acad. Sci. U S A. **104** (2007) 18760–18765.
37. Garrity, A.G., Pearlson, G.D., McKiernan, K., Lloyd, D., Kiehl, K.A., Calhoun, V.D.; Aberrant "default mode" functional connectivity in schizophrenia. Am. J. Psychiatry. **164** (2007) 450–457.
38. Calhoun, V.D., Kiehl, K.A., Pearlson, G.D.: Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. Hum. Brain Mapp. **29** (2008) 828–838.
39. Zhou, Y., Liang, M., Tian, L., Wang, K., Hao, Y., Liu, H., Liu, Z., Jiang, T.: Functional disintegration in paranoid schizophrenia using resting-state fMRI. Schizophr. Res. **97** (2007) 194–205.
40. Salvador, R., Martinez, A., Pomarol-Clotet, E., Sarro, S., Suckling, J., Bullmore, E.: Frequency based mutual information measures between clusters of brain regions in functional magnetic resonance imaging. Neuroimage. **35** (2007) 83–88.
41. Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E.: Neurophysiological architecture of functional magnetic resonance images of human brain. Cereb. Cortex. **15** (2005) 1332–1342.
42. Liang, M., Zhou, Y., Jiang, T., Liu, Z., Tian, L., Liu, H., Hao, Y.: Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. Neuroreport. **17** (2006) 209–213.
43. Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., Chen, Z., Zhu, C., He, Y.: Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. Hum. Brain Mapp., **30** (2009) 1511–1523.
44. Williamson, P.: Are anticorrelated networks in the brain relevant to schizophrenia?. Schizophr. Bull. **33** (2007) 994–1003.

Part II
Tamagawa-RIKEN Dynamic Brain Forum

Dynamical Systems and Accurate Temporal Information Transmission in Neural Networks

Alessandro E.P. Villa, Yoshiyuki Asai, Javier Iglesias, Olga K. Chibirova, Jérémie Cabessa, Pierre Dutoit, and Vladyslav Shaposhnyk

Abstract We simulated the activity of hierarchically organized spiking neural networks characterized by an initial developmental phase featuring cell death followed by spike timing dependent synaptic plasticity in presence of background noise. Upstream networks receiving spatiotemporally organized external inputs projected to downstream networks disconnected from external inputs. The observation of precise firing sequences, formed by recurrent patterns of spikes intervals above chance levels, suggested the build-up of an unsupervised connectivity able to sustain and preserve temporal information processing.

1 Introduction

The embryonic nervous system is initially driven by genetic programs that control neural stem cell proliferation, differentiation and migration through the actions of a limited set of trophic factors and guidance cues. After a relatively short period of stable synaptic density, a pruning process begins: synapses are constantly removed, yielding a marked decrease in synaptic density due to apoptosis – genetically programmed cell death – and selective axon pruning [1]. Overproduction of a critical mass of synapses in each cortical area may be essential for their parallel emergence through competitive interactions between extrinsic afferent projections [2]. Furthermore, background activity and selected patterns of afferent activity are likely to shape deeply the emergent circuit wiring [3]. Synapses can change their strength in response to the activity of both pre-, and post-synaptic cells following spike timing dependent plasticity (STDP) rules [4]. This property is assumed to be associated

A.E.P. Villa (✉)
Neuroheuristic Research Group, Grenoble Institute of Neuroscience, Université Joseph Fourier, Grenoble, France; Sleep Research Laboratory, Department of Psychiatry HUGE, Geneva, Switzerland; Neuroheuristic Research Group, ISI-HEC, University of Lausanne, Lausanne, Switzerland
e-mail: Alessandro.Villa@neuroheuristic.org

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_8,

with learning, synapse formation and pruning. In cell assemblies interconnected in this way some ordered and precise – in the order of few ms – interspike interval relationships, referred to as "spatio-temporal firing patterns" or "precise firing sequences", may recur within spike trains of individual neurons and across spike trains recorded from different neurons [5].

In this study we assume the existence of functional correlates of spatio-temporal neural coding, such that one would expect that the same temporal pattern of firing would be observed whenever the same information is processed in a network [6]. The relationship between the input and output spike trains is characterized by highly nonlinear transfer functions [7] and we investigate to which extent precise temporal information may be preserved assuming that each afference can carry only part of the overall information.

2 Methods

In the present study we simulate the activity of interconnected neural networks undergoing neural developmental phases. The output spike trains of the networks were scanned to detect precise firing sequences, simply referred below as "patterns", using the Pattern Grouping Algorithm (PGA) [8]. The structure and dynamics of the detected patterns were analyzed and compared with the results obtained for the single simulated networks in presence and in absence of stimuli. An underlying dynamical system attractor, if any, was searched following a technique of "denoising" the spike trains using the detected patterns [9]. The overall description of the simulation framework and parameters cannot be inserted here due to editorial space limitation and has been published elsewhere [10]. These characteristics naturally geared the modeling framework towards the analysis of spike trains recorded in a network of hierarchically organized neural networks.

3 Results

Appearance and disappearance of patterns was due to developmental changes shaped by STDP in the network connectivity underlying the process of temporally organized input activity. In absence of an external input more units survived at the end of the simulation run but less patterns were found in proportion to the number of active cells. Moreover, the ratio of detected patterns vs. active cells was larger in the downstream than in the upstream network. Figure 1 shows extreme cases of onset dynamics of a single-unit pattern: in one case a triplet appeared early in the network maturation and disappeared after $t \approx 35,000$ ms (Fig. 1a, b). The single-unit pattern <148C, 148C, 148C; 191±0.9, 698±1.0> was composed by spikes produced by unit #148C. This notation means that a precise firing sequence started with a spike of unit #148C, followed 191±0.9 ms later by a second spike of the same unit, and followed by a third spike 698±1.0 ms after the first. In the opposite case another pattern (a quadruplet in this example) appeared only at a later stage of maturation after

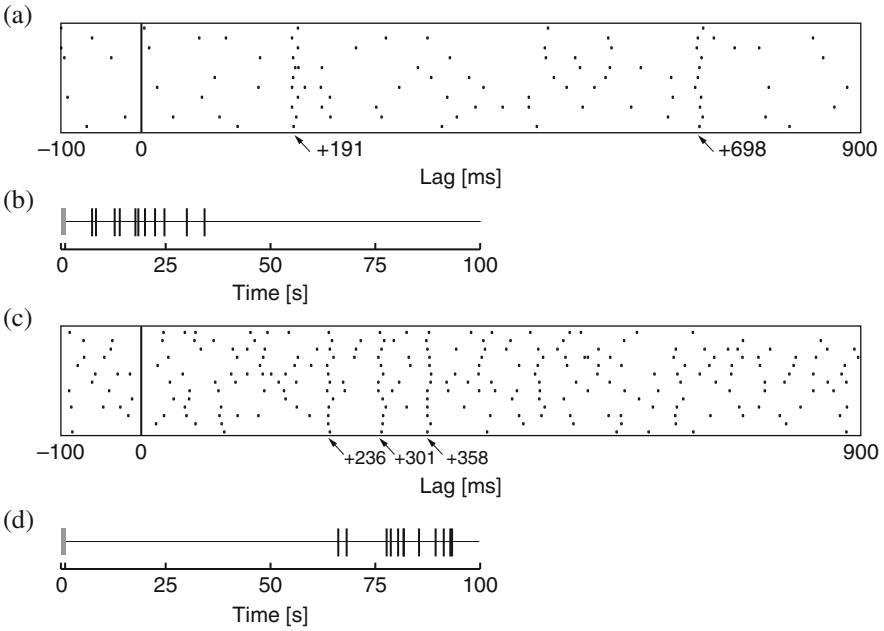


Fig. 1 Firing pattern < 148C, 148C, 148C; 191 ± 0.9, 698 ± 1.0 > that repeated 11 times (mean firing rate = 4.0 spikes/s). **a** Raster plot of the patterns aligned on the pattern start. **b** Raster plot of pattern onsets: each vertical tick corresponds to the onset time of each pattern occurrence; Firing pattern < 554, 554, 554, 554; 236±0.7, 301±0.8, 358±0.6 > (mean firing rate = 13.1 spikes/s). **c** Raster plot showing 13 repetitions; **d** Raster plot of pattern onsets

$t \approx 65,000$ ms (Fig. 1c, d). The nonlinear dynamic deterministic structure (attractor) embedded in the upstream afferent spike train was retrieved in the downstream spike train depending on the level of noise and also on parameters of the neuron model. We consider here the filtering effect produced by a cell assembly. Each neuron received only a fraction of a temporal information generated by a deterministic nonlinear dynamical system (Fig. 2). We used an input time series derived from the Zaslavskii map [9] and we observed that a distributed activity was much more efficient in transmitting the precise afferent temporal structure through the neural networks.

4 Discussion

Despite the fact that in downstream networks fewer cells were surviving at the end of the simulation run we found more firing patterns [11]. Downstream networks took more time to build-up the internal dynamics underlying the emergence of the patterns particularly in case of divergent and stronger connectivity towards the downstream network [12]. It is particularly interesting to note that when the external

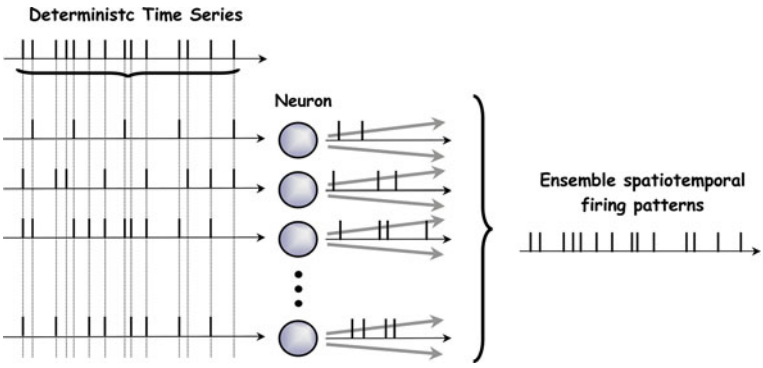


Fig. 2 The precise temporal information distributed in a cell assembly

input fed into the upstream network has en embedded complex temporally organized structure the temporally organized activity is distributed across the network and preserved throughout the downstream network and can be recovered by the analysis of reconstructed spike trains [13]. We suggest that specific neuronal dynamics characteristic of certain brain areas or associated to specific functional neuronal states play an essential role in the efficacy of transmitting a temporal pattern in a neuronal network. This feature is critical to determine the encoding and decoding processing that might be carried out by a single neuron and by the network and the extent of a distributed population coding scheme.

Acknowledgments The authors ackowledge the support by the EU FP6 grants #034632 PERPLEXUS and #043309 GABA.

References

1. Innocenti, G.M., Price, D.J.: Exuberance in the development of cortical networks. Nat. Rev. Neurosci. **6** (2005) 955–965.
2. Chechik, G., Meilijson, I., Ruppin, E.: Neuronal regulation: A mechanism for synaptic pruning during brain maturation. Neural. Comput. **11** (1999) 2061–2080.
3. Shatz, C.J.: Impulse activity and the patterning of connections during CNS development. Neuron **5** (1990) 745–756.
4. Roberts, P.D., Bell, C.C.: Spike timing dependent synaptic plasticity in biological systems. Biol. Cybern. **87** (2002) 392–403.
5. Abeles, M.: Corticonics: Neural Circuits of the Cerebral Cortex, 1st edn. New York, NY: Cambridge University Press (1991).
6. Villa, A.E.P.: Empirical evidence about temporal structure in multi-unit recordings. In Miller, R., ed.: Time and the Brain. Harwood Academic Publisher (2000) 1–51.
7. Segundo, J.P.: Nonlinear dynamics of point process systems and data. Int. J. Bifurcat. Chaos. **13** (2001) 2035–2116.
8. Tetko, I.V., Villa, A.E.P.: A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 1. detection of repeated patterns. J. Neurosci. Methods. **105** (2001) 1–14.

9. Asai, Y., Villa, A.E.P.: Reconstruction of underlying nonlinear deterministic dynamics embedded in noisy spike train. J. Biol. Phys. **34** (2008) 325–340.
10. Iglesias, J., Villa, A.: Emergence of preferred firing sequences in large spiking neural networks during simulated neuronal development. Int. J. Neural. Syst. **18** (2008) 267–277.
11. Chibirova, O.K., Iglesias, J., Shaposhnyk, V., Villa, A.E.P.: Dynamics of firing patterns in evolvable hierarchically organized neural networks. Lect. Notes. Comput. Sci. **5216** (2008) 296–307.
12. Iglesias, J., García-Ojalvo, J., Villa, A.E.P.: Effect of feedback strength in coupled spiking neural networks. Lect. Notes. Comput. Sci. **5164** (2008) 646–654.
13. Asai, Y., Guha, A., Villa, A.E.P.: Deterministic neural dynamics transmitted through neural networks. Neural. Netw. **21** (2008) 799–809.

Next Generation Large-Scale Chronically Implantable Precision Motorized Microdrive Arrays for Freely Behaving Animals

Jun Yamamoto

Abstract Multiple single-unit recording has become one of the most powerful in vivo electro-physiological techniques for investigating neural circuits. The demand for small and lightweight chronic recording devices with large number of electrodes has been increasing. However, conventional techniques do not allow fine adjustments to be made over large number of electrodes across multiple brain regions without disturbing the recording subject. Here, we have developed a large-scale motorized microdrive array to record from freely behaving animals. It has design flexibility and allows for remote adjustment of a large number of electrodes at a very fine resolution. Versions of the microdrive array were chronically implanted on both rats (21 microdrives) and mice (7 microdrives) and relatively long term recordings were taken. In this talk, current status of the new technology and future directions will be presented.

Keywords Motorized microdrive array · Large-scale · Chronically implantable microdrive · Freely behaving animals · Rats · Mice

1 Introduction

Electrophysiological recordings taken from freely behaving animals with chronically implanted multiple electrodes have played a critical role in our understanding of neuronal activity at the population level [1]. The recording devices used in these experiments employ various kinds of manually adjustable microdrives, which rely on miniature screws or threaded rods to advance electrodes [2, 3]. These microdrives must be small and light in weight, due to the physical limitations of the animal.

J. Yamamoto (✉)
Departments of Brain and Cognitive Sciences and Biology, RIKEN-MIT Neuroscience Research Center, The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
e-mail: yamajun@mit.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_9,

Here we sought to develop microdrive arrays for smaller animals that allow flexibility with regard to the number and location of the recording sites and the number of microdrives themselves, while minimizing electrical wire connections and overall weight. We developed a new type of motorized microdrive array that satisfies these requirements for both rats and mice. These microdrives have been extensively tested on multiple animals over the past four years. Long-term stable recording of multiple single units from multiple brain regions has been successfully achieved using these devices.

2 Methods

First, we briefly describe the key components of the electro-mechanical system used to achieve precision control in the large-scale microdrive array for freely behaving animals. (see Yamamoto and Wilson [4] for details)

2.1 Individual Motorized Microdrives

For one of the key devices of the project, we employed the smallest commercially available DC brushless servomotor with three stage 47:1 planetary gearbox (0206H05B+02/1 47:1, Faulhaber, Germany). Our motorized microdrive (Fig. 1a) has five basic components: the micromotor, threaded rod, slotted PEEK (polyetheretherketone) tube (PKT-062077, Small Parts, FL), shuttle lock, and a shuttle made with brass nuts. The entire microdrive weighs 300 mg, including 50 mg of magnetic shielding for the rat microdrive array. We designed two different scales of motorized microdrive arrays, one with twenty-one micromotors (weighing 29 g) for rats (Fig. 1b, c), and seven micromotors (weighing 4.8 g) for mice (Fig. 1d).

2.2 DC Servomotor Multiplexing Headstage

In order to manage dozens of motor wirings, we developed a lightweight High current capacity DC Servomotor Multiplexing Headstage with Serial Data Communication Technology (Fig. 2a). This headstage consists of numbers of multichannel analog switches (PI5V331Q, Pericom, CA) and multi-stage bus registers (CD4094B, Texas Instruments, TX). The analog switches are controlled by the outputs of multi-stage bus registers. The outputs of the bus register are encoded into serial bit stream. This technical solution allowed us to reduce numbers of motor connections to only nine wires (three analog sine waves plus six digital control lines including two power lines) regardless of the number of motors. The numbers of motors, which are supported by these headstages, are twenty-four and eight individual microdrives respectively.

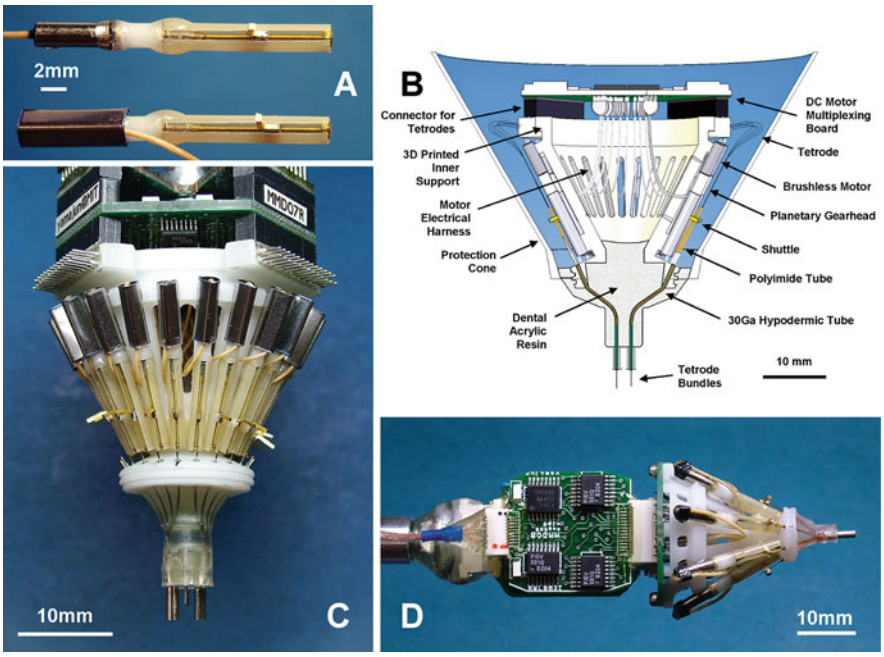

2mm
A
B
Connector for Tetrodes
3D Printed Inner Support
Motor Electrical Harness
Protection Cone
Dental Acrylic Resin
Tetrode Bundles
DC Motor Multiplexing Board
Tetrode
Brushless Motor
Planetary Gearhead
Shuttle
Polyimide Tube
30Ga Hypodermic Tube
10 mm
10mm
C
D
10mm

Fig. 1

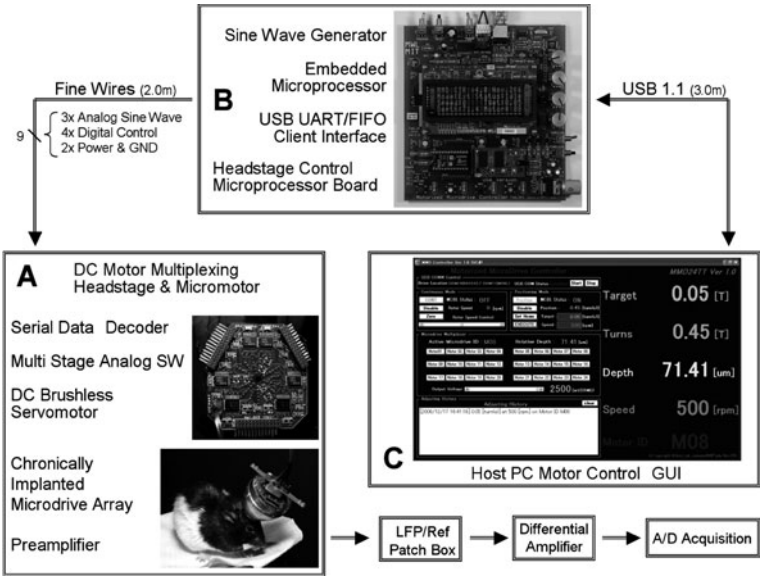

Sine Wave Generator
Embedded Microprocessor
B
USB UART/FIFO Client Interface
Headstage Control Microprocessor Board
Fine Wires (2.0m)
9
3x Analog Sine Wave
4x Digital Control
2x Power & GND
USB 1.1 (3.0m)
A
DC Motor Multiplexing Headstage & Micromotor
Serial Data Decoder
Multi Stage Analog SW
DC Brushless Servomotor
Chronically Implanted Microdrive Array
Preamplifier
C
Target 0.05 [T]
Turns 0.45 [T]
Depth 71.41 [um]
Speed 500 [rpm]
Host PC Motor Control GUI
LFP/Ref Patch Box
Differential Amplifier
A/D Acquisition

Fig. 2

2.3 Headstage Control Board and Control GUI Software

The headstage controller consists of three-channel sine wave generator (MCBL05002, Faulhaber, Germany) and embedded microprocessor (BS2SX, Parallax, CA) with USB UART/FIFO client interface LSI (FT2232C, Future Technologies Device International, UK) (Fig. 2b). The embedded microprocessor interfaces between the host computer and the sine wave generator through USB interface. It receives desired motor identification number, motor rotation speed and the amount of turn from the host computer program and returns current rotation speed and positional information to the host computer. The host PC controlling software is developed on Microsoft Visual C# 2005 (Microsoft, WA) that communicates the headstage control board through USB interface (Fig. 2c). The software has GUI based interface to control microdrives such as motor identification number, output power, rotation speed, direction and amount of turn etc. The information is updated every 250 ms through continuous communication between the software and the interface board.

3 Results

3.1 Chronically Implanted Microdrive Arrays

A total of seven male mice (30–35 g, Fig. 3a), and nine male rats (450–500 g, Fig. 3b) were tested. During these tests, two sets of rat microdrive arrays and another two sets of mouse microdrive arrays were used to test the electromechanical reliability as well as the reusability of the microdrive array. The duration of the experiments ranged from 2 to 5 months in both rat and mouse. On average, we took at least 2–3 weeks to reach desired recording target before we started neuronal recordings (example: mouse hippocampus CA1 area: Fig. 3c).

3.2 Fine Step Adjustments

One key advantage of the motorized microdrive array is ability to adjust electrodes at finer displacement steps to optimize unit isolation. To examine the potential effect

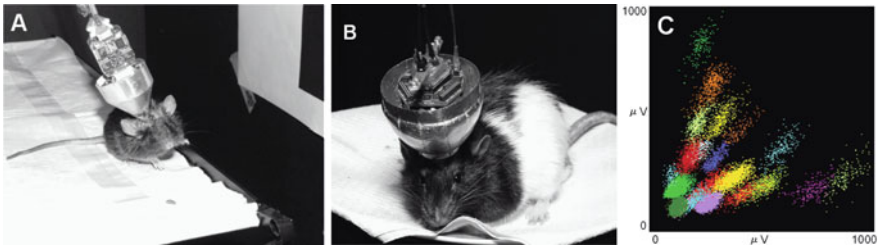


Fig. 3

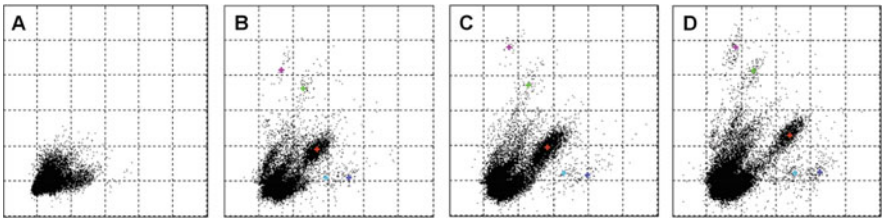


Fig. 4

of finer positional control, we tried to find a minimum adjusting step in which a systematic change could be seen in spike cluster distributions. Each panel in Fig. 4 illustrates the recording of spikes from multiple neurons on a single tetrode. Each data point represents a single spike event. These points are plotted at the peak amplitude of that spike on the first vs. the second channels of a four-channel tetrode. The figures (Fig. 4b–d) demonstrate that tetrodes advanced in small steps (8 μm/step) at the slowest speed (1.3 μm/s) showed systematic cluster movement. We also tried much finer displacement steps such as 4 or 2 μm, which had produced observable movement on the test bench. However, we could not confirm systematic cluster changes with these smaller displacements.

3.3 Stability Across Multiple Days

In addition to number of isolated units, the stability of isolated units across multiple days is an important issue. Figure 5 demonstrates a total of eleven individual units in rat somatosensory cortex that were stable across four consecutive days. These isolated units were obtained from deep layers (about 1,350 μm from the start) of the somatosensory cortex. Based upon the tests, we concluded that smaller steps and slower movements are essential factors for high quality recording. Overall, our experience suggests that the motorized microdrive array has advantages in the precision of the electrode positioning as well as isolation, yield, and stability of recorded units compared to conventional manual microdrive arrays.

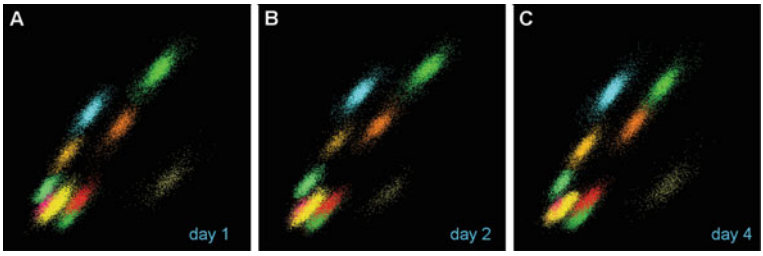


Fig. 5

4 Conclusion

Here, we have designed and tested two versions of large-scale motorized microdrive arrays for chronic recording in smaller animals, based on a modular and flexible design concept. The results demonstrated that our novel motorized microdrive array allowed us to obtain long-term recordings with quality comparable to, or exceeding that of conventional manual microdrive arrays [3]. The drive design is relatively simple and flexible allowing it to support recordings from brain structures at a variety of locations and depths. The major advantage of our new motorized microdrive array over existing techniques is the ability to remotely control a large number of microdrives slowly and precisely while maintaining the overall size and weight of conventional microdrive arrays [3]. Smaller steps and slower speed in the adjustments appear to enhance quality and stability of chronic recordings.

References

1. O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. **34** (1971) 171–175.
2. McNaughton, B.L., O'Keefe, J., Barnes, C.A.: The stereotrode: a new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. J. Neurosci. Methods. **8** (1983) 391–397.
3. Wilson, M.A., McNaughton, B.L.: Dynamics of the hippocampal ensemble code for space. Science. **261** (1993) 1055–1058.
4. Yamamoto, J., Wilson, M.A.: Large-scale chronically implantable precision motorized microdrive array for freely behaving animals. J. Neurophysiol. 100(4) (2008) 2430–2440.

Dynamic Receptive Fields in Auditory Cortex: Feature Selectivity and Organizational Principles

Christoph E. Schreiner and Craig A. Atencio

Abstract Spectrotemporal receptive fields (STRFs) capture the time-dependent behavior of spectral profile processing in auditory neurons. Here we review recent advances in utilizing these functional descriptors in exploring organizational principles of primary auditory cortex, such as cellular and laminar differences. In addition, we discuss an extension of the classical spectrotemporal filter approach that provides additional, independent spectrotemporal features that, combined with nonlinearities, allows a more complete characterization of cortical processing.

1 Modulation Differences Between Excitatory and Inhibitory Neurons

Excitatory pyramidal neurons and inhibitory interneurons have distinctive morphologic and electrophysiological properties and constitute the main elements of the cortical circuitry. The shape of action potentials can be used to differentiate between pyramidal neurons that often have longer action potentials and are classified as Regular-Spiking Units (RSUs), and chandelier and basket inhibitory neurons that have shorter action potentials and are classified as Fast-Spiking Units (FSUs). Potential physiological distinctions between these neuronal classes can be determined by stimulating primary auditory cortex neurons in the cat with a dynamic moving ripple stimulus and construct linear STRFs and their associated nonlinearities. Comparison of the two cell classes reveals distinct differences between the two cell classes in primary auditory cortex. FSUs generally have shorter latencies, greater stimulus selectivity, and higher temporal precision than RSUs. The STRF structure of FSUs is more separable than for RSUs, suggesting a relative independence between spectral and temporal processing aspects of the FSUs. The

C.E. Schreiner (✉)
Coleman Memorial Laboratory, W.M. Keck Foundation Center for Integrative Neuroscience, University of California, San Francisco, CA 94143-2206, USA
e-mail: chris@phy.ucsf.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_10,

nonlinearities associated with the two cell classes have a stronger asymmetry for FSUs, indicative of higher feature selectivity [1].

Temporal and spectral modulations are prominent aspects of communication sounds. Tuning to particular modulation frequencies is essential for correct speech identification, and speech can be correctly identified when only temporal modulation information and sparse carrier structure is retained. Likewise, spectral modulation information is important for communication sound processing since it is challenging for listeners to discriminate speech with a degraded spectral envelope. The importance of modulation processing prompted us to examine if modulation preferences operate as principles guiding receptive field properties. The finding that FSUs and RSUs receptive field share similar best temporal and spectral modulation frequency preferences indicates that they perform similar transformations of their thalamic modulation input for some properties, e.g., preferred modulation frequency, and different transformation for others, such as the shape of the MTFs. These distinctive functional differences between RSUs and FSUs provide a broader basis for unraveling the interactions between putative excitatory and inhibitory neurons that shape auditory cortical processing.

2 Laminar Organization of Spectral and Temporal Modulation Properties

Previous work in AI showed that modulation information can undergo a transformation between thalamus and cortex [2]. In comparison to thalamic output, neurons in the thalamocortical recipient layers 3b/4 follow slower modulations. Additionally, layers 3b/4 contain spatial topographies or local organization for characteristic frequency, latency, threshold, as well as spectral and binaural integration [3], [4]. However, after this initial stage of processing, little is known how the vertical AI microcircuit further shapes and transforms elemental acoustic information [5]. In particular, there is little evidence of functional layer-specificity despite clearly expressed cellular and connectional differences. Natural sounds are dominated by dynamic temporal and spectral modulations, and we used these properties to evaluate functional differences across cortical laminae. We examined the layer-specificity to acoustic modulation information by simultaneously recording from multiple AI laminae in the anesthetized cat. Neurons were challenged with dynamic moving ripple stimuli to compute spectrotemporal receptive fields (STRFs). From STRFs temporal and spectral modulation transfer functions (tMTFs, sMTFs) were calculated and compared across layers. Temporal and spectral modulation properties often differed between layers. On average, neurons in layer 2/3 and 6 respond to lower temporal modulations than those in layer 4. tMTFs were mainly bandpass in granular layer 4 and become more lowpass in infragranular layers. Compared to layer 4, spectral MTFs are broader and their upper modulation cut-off frequency are higher in layers 5, and 6. In individual penetrations, temporal modulation preference is similar across layers for ~70% of the penetrations, suggestive of a common,

columnar functional characteristic of temporal modulation preferences. By contrast, only ~30% of penetrations show consistent spectral modulation preferences across layers indicative of functional laminar diversity or specialization. AI layers that express differing modulation properties may sub-serve particular roles in the extraction of dynamic sound information specific to their targeted stations [6].

We have demonstrated that spectral and temporal modulation parameters, unlike CF, can vary significantly across AI layers. The behavior of temporal and spectral modulation processing is dissimilar in that temporal modulation has a stronger tendency for a columnar, layer-independent behavior. The direction of parameter changes is not tightly linked to a simple interlaminar flow pattern from thalamic input layers to supra- and infra-granular output layers. This constrains models of the laminar processing sequence of auditory cortex. Differences in modulation properties captured in STRFs make it more feasible to dissect laminar-specific, module-specific, and fieldspecific variations in the cortical processing regime and can help to determine whether common functional patterns pertain to cortical or subcortical inputs, and, combined with cell-type specific properties, reveal fundamental aspects of the local, lamina specific circuitry.

3 Joint Encoding of Multiple Auditory Features

STRFs, calculated through the spike-triggered average, have been used successfully to determine the modulation preferences and stimulus selectivity properties of auditory cortex neurons. While informative, STRFs may be biased by stimulus correlations and they do not characterize neural sensitivity to multiple stimulus dimensions. We explored ways to overcome these limitations by using a model in which a neuron is selective for two dimensions in a high dimensional stimulus space. To derive the model, single neuron responses were recorded in response to a dynamic moving ripple stimulus in the primary auditory cortex of the cat. Each relevant dimension was then reconstructed by finding those stimulus dimensions that maximized the mutual information between the neural response and a moving ripple stimulus [7]. This process removes the effects of stimulus correlations from the estimates of the dimensions. After the relevant dimensions were determined we calculated the nonlinear, memory-less input-output function that relates spiking probability to the stimulus projection onto the most informative dimensions (MIDs). For all analyzed neurons from cat primary auditory cortex we could establish two independent MIDs. For all neurons we found that the conventional spike-triggered averaged STRF was a good model of the first relevant dimension (MID1) but not for the other. The presence of a second spectrotemporal feature dimension (MID2) indicates that AI neurons are sensitive to multiple stimulus dimensions. For the majority of neurons the nonlinear input-output function for the first MID generally was asymmetric and often monotonic in shape. By contrast, the symmetric for the MID2 nonlinearity was symmetric and nonmonotonic. We determined the cooperativity between the first and second dimensions in spike-generation by mapping

the two-dimensional input-output function. By constructing a joint two-dimensional nonlinearity we found that the interactions between the two MIDs were synergistic (jointly accounting for a higher mutual information than the sum of the two MIDs) indicating that the relevant dimensions must be considered together, as opposed to independently estimating receptive field models. These results indicate that this approach promises to be useful in primary auditory cortex and in secondary areas for which complex sounds may be more appropriate stimuli. The striking difference between the nonlinearities of the two MIDs suggests the concurrent and cooperative implementation of different tasks by each neuron, such as feature detection and context evaluation.

Analysis of the contribution of each feature filter to the stimulus information in the spike train showed that the relative contributions fell along a continuum. An implication from this is that for some neurons the first dimension, or even the spiketriggered average, is a sufficient approximation to the overall processing taking place. For most neurons, however, a greater diversity of computation is evident. For these neurons one feature dimension is not an adequate description and the standard spiketriggered average model is not sufficient. Coupled with the understanding that the multiple stimulus dimensions of many AI neurons operate synergistically it becomes clear that an extended linear-nonlinear model allows us to recover a much more complete picture of the neuronal coding strategies employed by AI neurons. The application of the linear-nonlinear model to AI using information maximization represents a significant advance over previous approaches, since it permits quantification of several stimulus features that influence a neuron's firing, how feature selective a neuron is, and how this feature selectivity interacts synergistically to influence a neuron's response.

A functional interpretation of the components of the two-filter model remains speculative, especially in light of the strong interaction between the two nonlinearities. Based on insights from the simple and complex cells in the visual system it is not unreasonable to postulate that the first filter in combination with the asymmetric nonlinearity is a feature or foreground-structure detector tuned to a narrow range of stimulus constellations. The closer the match between stimulus and filter, the stronger is the response. By contrast, the second filter with its associated symmetric nonlinearity more likely corresponds to an envelope-phase insensitive feature detector tuned to a broad range of stimulus constellations. The main purpose of the second filter may be a gain control for feature sensitivity in a context-dependent manner to improve the relevant auditory foreground features in various acoustic environments.

4 Conclusions

Auditory processing by cortical neurons shows specificity to the cell types (excitatory/inhibitory) as well as to the laminar position of the cells in the columnar circuit. Single- or multi-feature linear spectrotemporal receptive fields allow a more complete description than most other approaches of the stimulus-based response properties of central auditory neurons. The expandability of this approach to

multiple feature filters with the addition of associated nonlinearities provides clear advantages in distinguishing different neuronal classes.

Acknowledgments Supported by NIH grant DC02260.

References

1. Atencio, C.A., Schreiner, C.E.: Spectrotemporal processing differences between auditory cortical fastspiking and regular-spiking neurons. J. Neurosci. **28** (2008) 3897–3910.
2. Miller, L.M., Escabí, M.A., Read, H.L., Schreiner, C.E.: Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. J. Neurophysiol. **87** (2002) 516–527.
3. Middlebrooks, J.C., Dykes, R.W., Merzenich, M.M.: Binaural response-specific bands in primary auditory cortex (AI) of the cat: topographical organization orthogonal to isofrequency contours. Brain Res. **181** (1980, 1981) 31–48.
4. Schreiner, C.E., Read, H.L., Sutter, M.L.: Modular organization of frequency integration in primary auditory cortex. Annu. Rev. Neurosci. **23** (2000) 501–529.
5. Linden, J.F., Schreiner, C.E.: Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. Cereb. Cortex. **13** (2003) 83–89.
6. Atencio, C.A., Schreiner, C.E.: Laminar diversity of dynamic sound processing in cat auditory cortex. J. Neurophysiol. **103** (2009, 2010) 195–205.
7. Atencio, C.A., Sharpee, T.O., Schreiner, C.E.: Cooperative nonlinearities in auditory cortical neurons. Neuron. **58** (2008) 956–966.

Top-Down Mechanism of Perception: A Scenario on the Role for Layer 1 and 2/3 Projections Viewed from Dynamical Systems Theory

Hiroshi Fujii, Kazuyuki Aihara, and Ichiro Tsuda

Abstract Cortical layer 1 is the main entrance of top-down signals from other higher cortices and subcortical nuclei. Recent findings challenge the view that top-down signals play just a modulatory role. However, how top-down signals are implemented to help reconstruct the internal representations of images, or an episode of events is poorly understood. Recent experimental data suggest that the "ongoing" brain state without *external* inputs into layer 4 and attentional top-down signals continually fluctuates among the intrinsic patterns of activity. Viewed from a dynamical systems standpoint, the transitory intrinsic states could be an expression of "attractor ruins", observed in a *mesoscopic* dynamical system. We hypothesize that when top-down signals arrive as attention(s), contexts, or indices, the local brain state is temporarily stabilized to be an attractor(-like) state. In view of the anatomical and physiological configurations of neuronal systems in the superficial layers, we propose a hypothetical scenario for such non-classical dynamics in the brain.

Keywords Cortical layer 1 · Superficial layers · Top-down mechanism of perception · Internal representations of images · Attentional top-down signal · Attractor ruin · Milnor attractor · Presynaptic inhibition of GABA release · Muscarinic acetylcholine · Glutamatergic spike volley · Parvalubmin-positive fast spiking neuron · Nicotinic depolarization · Carletinin-positive interneuron

1 Cortical Layer 1: The Crowning Mystery

Cortical layer 1 s receive projections from a number of other cortical areas together with the subcortical nuclei, the thalamic matrix elements (including pulvinar), and the hippocampal CA1 region (via the enthorhinal cortex). As Hubel noted in 1982

H. Fujii (✉)
Department of Intelligent Systems, Kyoto Sangyo University, Kyoto 603-8555, Japan;
Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan
e-mail: fujii@bacco.kyoto-su.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_11,

[1], layer 1 provides the crowning mystery in understanding neocortical functions: in turn understanding layer 1 functions may provide a clue for resolving questions related to cortical organization and dynamics.

One thing has long been established about layer 1: it is where top-down signals enter as attention, contexts, and modulating inputs. This fact is in sharp contrast with layer 4, which specializes in receiving specific bottom-up signals from the thalamic core elements and from layer 2/3 of the other lower cortical areas [2].

The following is the typical, classical statement about the role of top-down projections onto layer 1: "Feedback projections (via layer 1) have modulatory effects while activity is mainly driven by the bottom-up pathway via layer 4." However, recent experimental data pose difficulties in understanding layer 1 functions along such a standpoint. Two counter-examples include the recent experimental observation of Kenet et al. [3], and "mental imagery" [4].

An implication of the Kenet et al. data is that, a number of local *internal states* are built-in with hierarchical structures in the cortical circuits. The 'ongoing' brain state in the absence of *external* inputs into layer 4 is not random, but continually fluctuating among those intrinsic patterns of activity if, because of anesthesia, *no* or at least *reduced* attentional top-down signals are projected onto layer 1.

The mental imagery poses a further question: how top-down signals are implemented to help reconstruct internal representations of images, or an episode of recent events[1] in the absence of external stimuli?

1.1 Top-Down Projections: Two Concurrent Flows

Associated with top-down *attentions,* whether overt or covert [5], or even in mental imagery [6], we may postulate that two concurrent flows are projected onto the cortex. The first one is glutamatergic (Glu) spike volleys projecting onto layer 1 from "higher" cortical areas and the thalamic matrix elements, and the second is cholinergic (and GABAergic) projections on all six layers from the nucleus basalis of Meinert (NBM), which are triggered by Glu spikes from the medial prefrontal cortex [7]. These two concurrent flows, which we call simply as "attentional flows", are the main players of the *top-down mechanism*.

2 A Dynamical Systems-Theoretic Scenario

How can the brain dynamics described as "continually fluctuating among intrinsic patterns of activity" be characterized from a dynamical systems standpoint?

[1] In "mental imagery", no external stimuli play a role. When one is asked a question such as "which is longer – a donkey's ears or an ear of corn?", one may visualize the objects, and "see" the necessary properties, as most people report [4]. As Kosslyn pointed out, this sort of introspection suggests that visual mental images reconstruct the spatial geometry of objects [4].

The first key is that an "intrinsic state" could be regarded as an "attractor ruin" in the dynamical systems term, which Tsuda first formulated in [8]. An attractor ruin must have a mechanism of allowing both transition and reinjection to and from a state. A typical example of an attractor ruin is a perturbed structure of the non-classical Milnor attractor [9], which possesses a positive measure of attracting orbits, but simultaneously may possess repelling orbits from itself.[2]

The second key we propose is the notion that inhibition may destabilize attractors under certain conditions [10]. Default inhibitions on potentially existing attractors may transform them into fragmentary "ruins". Such a destabilizing mechanism has been described by theoretical and numerical works with a variety of contexts [8, 11, 12]. In other words, we claim that, in the absence of external stimuli the internal states potentially exist only in the form of attractor ruins before the arrival of attentional flows, and hence the cortical dynamics exhibits continual transitions among ruins. This may characterize the *default* cortical dynamics. We hypothesize that attentional flows temporarily reverse this process, and restore an attractor state making ruins as its building blocks.

How do attentional flows make such state transitions of cortical dynamics? We propose a hypothetical scenario in view of the anatomical and physiological configurations of the neuronal system in the superficial layers.

2.1 FS γ Hypothesis

We postulate that the *local* emergence of the gamma band rhythm is carried by layer 2/3 parvalbumin-positive (PV^+) fast spiking (FS) neurons, driven by Glu spikes from layer 4 (and, possibly layer 6 [13]). The FS γ rhythm may lightly entrain layer 2/3 pyramidal neurons (PYRs). See, Cardin et al., 2009 [14]. The FS γ rhythm may provide a narrow temporal window to PYRs for effective excitation, resulting in the *local* synchrony of neuronal ensembles [14].

We postpone discussion of the global γ synchrony over multiple cortices as the prerequisite for perception until the last paragraph.

2.2 Attentional ACh Releases PYRs from the Ongoing Inhibitions by the Presynaptic Disinhibition: Its Consequence

In layer 2/3 both PYRs and PV^+ FS cells are *unresponsive* postsynaptically to acetylcholine (ACh) [15]. This poses the question: what kind of roles may ACh projections in fact carry onto layer 2/3?

[2]The concept of attractor ruins may include a wider class of non-classical attractors than the Milnor attractor. By this reason, we may use the term "attractor ruins" in this paper to include possible but unknown classes of ruins.

Evidence exists that ACh release onto layer 2/3 may function in a presynaptic mechanism with muscarinic ACh receptor (mAChR) [16]. Release of ACh in layer 2/3 activates the m2 subtype of mAChR at synapses of $GABA^+$ INs, which presynaptically inhibits $GABA_A$ release onto PYRs. (See, Fig. S5 in Supplement of [16].) This dynamical strengthening of network connectivity transforms *attractor landscape*[3] from Milnor-like attractor ruins to a "relevant" *stable* attractor(-like) state making ruins as its building blocks [10, 11, 17, 18].

However, collaboration with the second top-down flow, i.e., Glu spike volleys projected onto layer 1 might be essential making the binding of fragmentary ruins.

2.3 Top-Down Glu Spike Volleys in Layer 1 – As Indexing Signals for Designating and Binding "Internal States"

The top-down Glu spike volleys project on distal apical dendrites of layer 2/3 (and layer 5) PYRs at layer 1. We hypothesize that, at least some of, these Glu spike volleys function as *indexing signals* for designating and locally binding the internal states to form hierarchically organized attractor states.

Two remarks follow. First, such a view on layer 1 projections has been repeatedly expressed in the literature with various contexts. (See, e.g., for "index theory" [19], or "tag" [20].) The second remark concerns dynamic pattern completion, or chaotic associative memory. As Nara et al. reported [12, 21], a *strike* of spike volleys as a memory fragment was enough to recover the corresponding attractor state, and to move the orbit to the vicinity of the assigned attractor in their chaotically transitory networks [21]. Note that such a property is commonly observed and characterizes dynamics itinerant among attractor ruins. (See, also Freeman [22].)

2.4 Top-Down Mechanism of Perception Viewed from Dynamical Systems Theory

The role of ACh projection onto layer 1 constitutes one of the main mysteries proposed by Hubel [1]. It works only via a nicotinic ACh receptor (nAChR) activation, and that produces a rapid depolarization of layer 1 calretinin-positive (CR^+) interneurons (INs) [23]. The main targets of the CR^+ IN axons in layer 2/3 are $GABA^+$ INs [23], but the exact subtypes of target $GABA^+$ INs remain unclarified. As suggested in Gabott [24], an initial candidate could be PV^+ FS cells that inhibit PYRs peri-somatically receiving feed forward excitations from layer 4. Another possibility is some other PV^- INs, that are probably fed by PYRs, and in turn inhibit PYRs at distal apical dendrites. The two possibilities are not exclusive.

[3] "Attractor landscape" is usually used for potential systems. We may use the term here to mean the spatial structure of basins of attractors.

Focusing on the PV^+ FS cell target story, we may have two alternatives. The inhibitory actions on layer 2/3 PV^+ FS cells due to rapid nicotinic depolarization of layer 1 CR^+ INs bring about instantaneous inhibitions of PV^+ FS cells. In view of its *global* nature, top-down attentions with cholinergic projections via NBM might function for *resetting* the phase and *starting up* the synchrony in the gamma band over multiple cortices. Alternatively, this hypothetical phase resetting might be done within local areas. In such a case, the interareal phase synchrony could be a result of entrainment by the hippocampal theta rhythm [25].

3 Concluding Remarks

The aim of this paper was, inspired by recent experimental observations of cortical dynamics, to draw a possible theoretical and hypothetical picture of the top-down mechanism for perception based on the physiological and anatomical data with the emphasis on dynamical systems viewpoints. We hope that it may motivate further studies on this subject from both experimental and theoretical standpoints.

Acknowledgements The first author (HF) was supported by a Grant-in-Aid for Scientific Research (C), No. 19500259, from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government.

References

1. Hubel, D.H.: Cortical neurobiology: a slanted historical perspective. Annu. Rev. Neurosci. **5** (1982) 363–370.
2. Felleman, D.J., Van Essen, D.C.: Distributed hierarchical processing in primate cerebral cortex. Cereb. Cortex. **1** (1991) 1–47.
3. Kenet, T., et al.: Spontaneously emerging cortical representations of visual attributes. Nature. **425** (2003) 954–956.
4. Kosslyn, S.M.: Mental images and the brain. Cogn. Neuropsychol. **22** (2005) 333–347.
5. Thompson, K.G., et al.: Neural basis of covert attention in the frontal eye field. J. Neurosci. **25** (2005) 9479–9487.
6. Ishai, A., et al.: Visual imagery of famous faces: effects of memory and attention revealed by fMRI. NeuroImage. **17** (2002) 1729–1741.
7. Golmayo, L., et al.: Electrophysiological evidence for the existence of a posterior cortical–prefrontal–basal forebrain circuitry in modulating sensory responses in visual and somatosensory rat cortical areas. Neuroscience. **119** (2003) 597–609.
8. Tsuda, I., Koerner, E., Shimizu, H.: Memory dynamiacs in asynchronous networks. Prog. Theor. Phys. **78** (1987) 51–71.
9. Milnor, J.: On the concept of attractor. Commun. Math. Phys. **99** (1985) 177–195.
10. Tsuda, I.: Chaotic itinerancy as a dynamical basis of hermeneutics of brain and mind. World Futures. **32** (1991) 167–185.
11. Fujii, H., et al.: Functional relevance of "excitatory" GABA actions in cortical interneurons: a dynamical systems approach. J. Integr. Neurosci. **3** (2004) 183–205.
12. Nara, S., et al.: Chaotic memory dynamics in a reccurrent neural networks with cyclic memories embedded by pseudo-inverse method. Int. J. Bifurcat. Chaos. **5** (1995) 1205–1212.
13. Briggs, F., Callaway, E.M.: Layer-specific input to distinct cell types in layer 6 of monkey primary visual cortex. J. Neurosci. **21** (2001) 3600–3608.

14. Cardin, J.A.: Driving fast-spiking cells induces γ-rhythm and controls sensory responses. Nature. **459** (2009) 663–668.
15. Gulledge, A.T. et al.: Heterogeneity of phasic cholinergic signaling in neocortical neurons, J. Neurophysiol. **97** (2007) 2215–2229.
16. Kruglikov, I., Rudy, B.: Perisomatic GABA release and thalamocortical integration onto neocortical excitatory cells are regulated by neuromodulators. Neuron. **58** (2008) 911–924; Online Supplement.
17. Adachi, K., Aihara, K.: Associative dynamics in chaotic neural network. Neural Netw. **10** (1997) 83–98.
18. Tsuda, I.: Dynamic link of memories—chaotic memory map in non-equilibrium neural networks. Neural Netwo. **5** (1992) 313–3226.
19. Teyler, T.J., et al.: The hippocampal indexing theory and episodic memory: updating the index. Hippocampus. **17** (2007) 1158–1169.
20. Bar, M.: A Cortical Mechanism for Triggering Top-down Facilitation in Visual Object Recognition. J. Cogn. Neurosci. **15** (2003) 600–609.
21. Kuroiwa, J., et al.: Sensitive response of a chaotic wandering state to memory fragment inputs in a chaotic neural network model. Int. J. Bifurcat. Chaos. **14** (2004) 1413–1421.
22. Freeman, W.J.: Tutorial on neurobiology: from single neurons to brain chaos. Int. J. Bifurcat. Chaos. **2** (1992) 451–482.
23. Christophe, E., et al.: Two types of nicotinic receptors mediate an excitation of neocortical layer I interneurons. J. Neurophysiol. **88** (2002) 1318–1327.
24. Gabott, P.L.A., Hacon, S.J.: Local circuit neurons in the medial prefrontal cortex (Areas 24a,b,c, 25 and 32) in the Monkey: I. Cell Morphology and Morphometrics. J. Comp. Neurol. **364** (1996) 567–608.
25. Sirota, A. et al.: Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. Neuron. **60** (2008) 683–697.
26. Lee, M.K.: Cholinergic basal forebrain neurons burst with thetaa during waking and paradoxical sleep. J. Neurosci. **25** (2005) 4365–4369.

Beyond Sensory Coding: The Cognitive Context of Olfactory Neurodynamics

Leslie M. Kay

Abstract The local field potential from the mammalian olfactory system displays several oscillatory modes. Gamma oscillations (40–100 Hz) are associated with odor stimulation in the olfactory bulb (OB) and are enhanced with odor discrimination learning in a 2-alternative choice task when odors are very similar. Beta oscillations (15–30 Hz) are associated with learning in a go/no-go task. These two oscillatory modes represent different cognitive states, and different networks appear to support them. Gamma oscillations are produced locally within the OB, while beta oscillations rely on bidirectional connections with the rest of the olfactory and limbic systems. These results help elucidate the context-dependent sensory representation seen at every neurophysiological level in the OB.

Keywords Gamma oscillations · Beta oscillations · Operant behavior · Olfaction

1 The Primary Olfactory System: Anatomy and Physiology

The mammalian olfactory system displays a rich array of oscillations, which can be seen in the cortical local field potential (LFP). These signals represent the summed synaptic activity from many neurons (100s to 1,000s). When combined with behavioral tasks, analysis of the LFP signal from multiple brain regions produces a dynamic picture of cognitive processing within early sensory areas normally assumed to be objective feature detectors. Neurons that receive sensory input from the nasal epithelium respond in complex ways to odors, and their odor selectivity changes as stimulus associations change [1–3]. In addition, the amplitude pattern of a common LFP waveform over the olfactory bulb (OB) depends strongly on experience [4].

L.M. Kay (✉)
Department of Psychology, Institute for Mind and Biology, The University of Chicago, Chicago, IL 60637, USA
e-mail: lkay@uchicago.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_12,

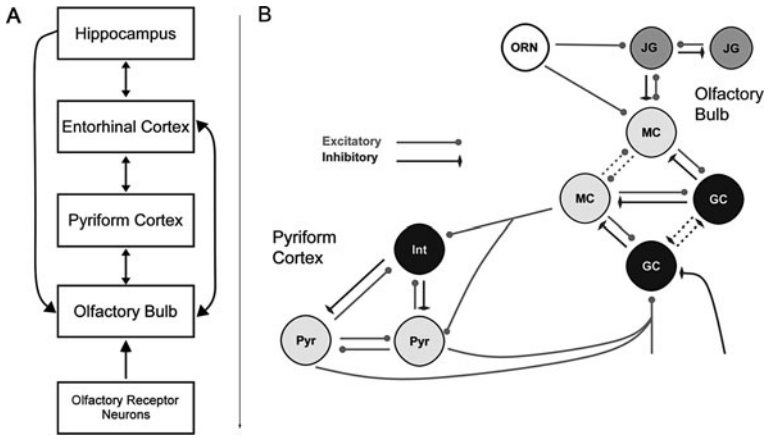


Fig. 1 **a** Schematic of the olfactory system and connections with the hippocampal system. The olfactory bulb receives input from all parts of this system. **b** Sensory input from olfactory receptor neurons *(ORN)* reaches mitral cells *(MC)* in glomeruli at the OB periphery. Juxta-glomerular cells *(JG)* modify the input signal. Excitatory MCs make reciprocal synapses with inhibitory granule cells *(GC)*. There are also inhibitory and excitatory inputs to GCs from other brain areas. Pyriform cortex neurons receive inputs from the OB onto excitatory pyramidal cells *(Pyr)* and inhibitory interneurons *(Int)*. Pyramidal cells project back to OB GCs

Olfactory sensory receptors in the nasal epithelium project directly to the cortical olfactory bulb (OB). The neurons that receive this input project to many parts of the olfactory and limbic systems (Fig. 1). Within the OB, excitatory and inhibitory connections serve to sharpen or otherwise modify the signal (Fig. 1b).

Three characteristic oscillations have been associated with specific behaviors (reviewed in [5]). Theta oscillations (2–12 Hz; Fig. 2a) follow the respiratory drive. Olfactory bulb (OB) gamma oscillations (40–100 Hz; Fig. 2a, c), first described by Adrian, have since been examined by scores of researchers [6]. These oscillations initiate at the transition from inhalation to exhalation and have been associated with odor stimulation. Gamma oscillations are further subdivided into gamma1 (60–100 Hz; Fig. 2a, c), the sensory-evoked oscillations, and gamma2 (~50–60 Hz; Fig. 2b), which are associated with alert states and occur between breaths during periods of slow respiratory rates (< 5 Hz). Beta oscillations (15–30 Hz; Fig. 2d) are produced during odor stimulation in some tasks when a rat has learned to discriminate odors.

2 Two Behavioral Tasks and Two Oscillatory Modes

The behavioral context in which an individual samples an odor can strongly affect perception [7], the spatial pattern of the olfactory EEG [4, 8], and neuronal responses to odors [1–3]. Thus, it is reasonable to assume that in addition to the associative context of an odor (positive vs. negative reinforcement), the cognitive structure of the task itself should have consequences on the neural representations in

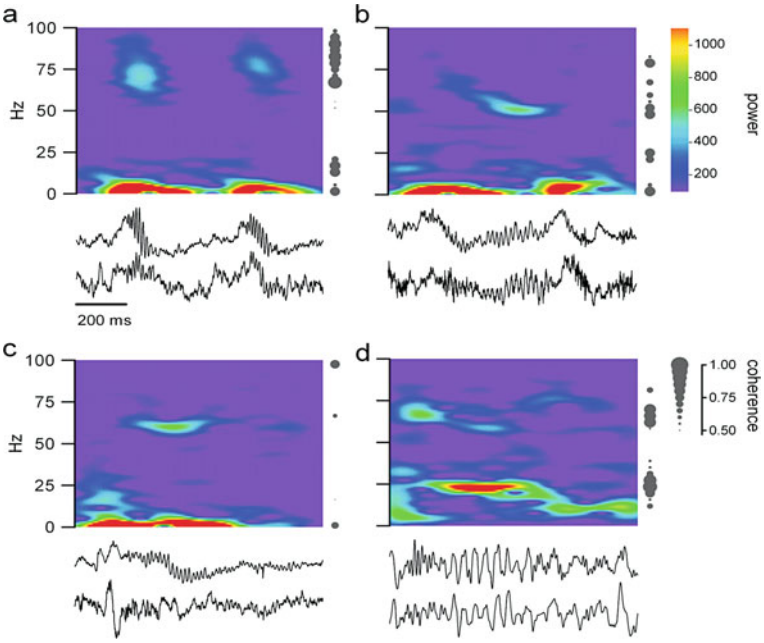


Fig. 2 LFP oscillations in OB and aPC. **a** Gamma1 initiates at the inhalation peak. The *lower frequency wave* shows two inhalation cycles. Gamma1 and theta coherence *(gray dots to right)* is strong. **b** Gamma2 during alert immobility is seen in both structures between inhalations. **c** Gamma1 is strong in the OB and absent in the aPC during high overlap odor discrimination in a 2AC task. **d** beta during odor exposure in a GNG task. (Figure from Kay et al. [5], with permission)

the OB and other parts of the olfactory system. It is known that rodents can identify odors very quickly, but their performance and the amount of time they spend sampling the odor is influenced by the difficulty of the discrimination (amount of overlap of the input patterns) and the type of task they use to make the discrimination, respectively [9].

We used two different tasks to evaluate discrimination ability and neurophysiological networks in rats: the Go/No-Go (GNG) and 2-alternative choice (2AC) tasks. In the GNG task, the rat responds to one odor with a lever press and refrains from pressing the lever for the other odor. In the 2AC task, there are two response levers, and the rat presses one lever in response to one odor and the other in response to the second odor. Correct lever presses are rewarded with a sucrose pellet, and incorrect lever presses result in an immediate end to the trial (lights out) and an extra delay before the next trial. In our laboratory, rats easily learn the GNG task and new odor sets, responding with >90% accuracy [10, 11]. The 2AC task is more difficult to learn and transfer to new odors requires longer [12].

In the 2AC task, when odorants are very different there is no enhancement of beta or gamma1 with increases in performance, but when the odorants are very similar, there is a significant increase in gamma1 for both odors (Fig. 2c; [12]). In the GNG task, gamma1 oscillations are depressed, and when the rats reach criterion

performance oscillations in the beta band are significantly enhanced (Fig. 2d; [11, 13]). There is no increase in beta power during discrimination of highly overlapping input patterns. Beta oscillations are present in the 2AC task, but beta power is not related to performance. This task-dependence of the oscillatory mode is not just a difference in frequency. During the GNG task, beta oscillations show strong coherence with beta oscillations in the aPC (Fig. 2d, gray dots to the right of the graph), but during the 2AC task, gamma1 oscillations show almost no coherence with gamma band activity in the aPC (Fig. 2c). Previous studies have shown that decreasing feedback to the OB from other brain regions increases gamma oscillations and cooperative firing patterns among MCs, but this same manipulation destroys the beta oscillation mode [14]. These data suggest that different brain networks support gamma1 and beta oscillatory modes.

3 Cognitive and Possible Circuit Differences

Differences in oscillatory modes dependent on the cognitive task may not be as surprising when psychological studies are taken into consideration. GNG tasks are generally considered to be easier to learn, but these tasks also require inhibiting a learned, almost rote, motor response. Response inhibition requires circuits including the basal ganglia and frontal cortex [15]. 2AC tasks also require some form of inhibition, as a common strategy for a rat is to favor one lever over the other for a large block of trials, performing at near 100% for one odor and at much lower levels for the other odor. However, the overwhelming cognitive mode in the 2AC task is a binary decision, which likely involves either different brain regions or the same regions in a different mode than for the GNG task [16]. Future studies should uncover both common and divergent networks associated with these two tasks, and these networks should differentially influence the type and degree of oscillatory activity via feedback mechanisms to the OB and neuronal firing patterns.

Acknowledgments Supported by NIDCD R01DC007995.

References

1. Kay, L.M., Laurent, G.: Odor- and context-dependent modulation of mitral cell activity in behaving rats. Nat. Neurosci. **2**(11) (1999) 1003–1009.
2. Rinberg, D., Koulakov, A., Gelperin, A.: Sparse odor coding in awake behaving mice. J. Neurosci. **26**(34) (2006) 8857–8865.
3. Doucette, W., Restrepo, D.: Profound context-dependent plasticity of mitral cell responses in olfactory bulb. PLoS Biol. **6**(10) (2008) e258.
4. Freeman, W.J., Schneider, W.: Changes in spatial patterns of rabbit olfactory EEG with conditioning to odors. Psychophysiology. **19**(1) (1982) 44–56.
5. Kay, L.M., Beshel, J., Brea, J., Martin, C., Rojas-Líbano, D., Kopell, N.: Olfactory oscillations: the what, how and what for. Trends Neurosci. **32**(4) (2009) 207–214.
6. Rojas-Líbano, D., Kay, L.M.: Olfactory system gamma oscillations: the physiological dissection of a cognitive neural system. Cogn. Neurodyn. **2**(3) (2008) 179–194.

7. Herz, R.S.: The effect of verbal context on olfactory perception. J. Exp. Psychol. Gen. **132**(4) (2003) 595–606.
8. Barrie, J.M., Freeman, W.J., Lenhart, M.D.: Spatiotemporal analysis of prepyriform, visual, auditory, and somesthetic surface EEGs in trained rabbits. J. Neurophysiol. **76**(1) (1996) 520–539.
9. Rinberg, D., Koulakov, A., Gelperin, A.: Speed-accuracy tradeoff in olfaction. Neuron. **51**(3) (2006) 351–358.
10. Frederick, D.E., Barlas, L., Ievins, A., Kay, L.M.: A critical test of the overlap hypothesis for odor mixture perception. Behav. Neurosci. **123**(2) (2009) 430–437.
11. Martin, C., Beshel, J., Kay, L.M.: An olfacto-hippocampal network is dynamically involved in odor-discrimination learning. J. Neurophysiol. **98**(4) (2007) 2196–2205.
12. Beshel, J., Kopell, N., Kay, L.M.: Olfactory bulb gamma oscillations are enhanced with task demands. J. Neurosci. **27**(31) (2007) 8358–8365.
13. Martin, C., Gervais, R., Hugues, E., Messaoudi, B., Ravel, N.: Learning modulation of odor-induced oscillatory responses in the rat olfactory bulb: a correlate of odor recognition?. J. Neurosci. **24**(2) (2004) 389–397.
14. Martin, C., Gervais, R., Messaoudi, B., Ravel, N.: Learning-induced oscillatory activities correlated to odour recognition: a network activity. Eur. J. Neurosci. **23** (2006) 1801–1810.
15. Aron, A.R., Durston, S., Eagle, D.M., Logan, G.D., Stinear, C.M., Stuphorn, V.: Converging evidence for a Fronto-Basal-Ganglia network for inhibitory control of action and cognition. J. Neurosci. **27**(44) (2007) 11860–11864.
16. Braver, T.S., Barch, D.M., Gray, J.R., Molfese, D.L., Snyder, A.: Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cereb. Cortex. **11**(9) (2001) 825–836.

Temporo-Parietal Network Model for 3D Mental Rotation

Toshio Inui and Mitsuru Ashizawa

Abstract The ability to rotate objects mentally has been suggested to be related to recognition of visual objects presented from non-canonical viewpoints. However, the neural mechanism underlying this ability is still unclear. In this paper, a global neural network model is proposed. This model consists of two subsystems, a parietal network for mental rotation, and inferior temporal network for object recognition. In this model, it is assumed that mental rotation is realized by a process in which the egocentric representation of objects in the intraparietal sulcus is rotated by motor signals that are internally generated in the premotor cortex. The rotated information is sent downward to the visual cortex as a rotated visual image; meanwhile, object recognition is achieved by a matching process with target object images in the inferior temporal cortex. The parallel distributed processing of this model achieves robust object recognition from various viewpoints including the non-canonical view.

Keywords Mental rotation · Object recognition · Neural network model

1 Introduction

It has been reported that right parietal cortex is involved in the object recognition from the non-canonical view [1, 2]. Schendan and Stern [3] identified the brain network for object recognition according to an object recognition theory called "multiple view plus transformation (MVPT)" [4].

They found that the left occipital and temporal cortex were activated for object categorization, whereas the CIP (caudal intra-parietal), TOS (transverse occipital sulcus), and right posterior ventro-lateral prefrontal cortex (BA44/6) were activated for mental rotation.

T. Inui (✉)
Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
e-mail: inui@i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_13,

According to Kveraga et al. [5], for the object recognition, the candidate of categories would be extracted in the left posterior ventro-lateral prefrontal cortex through the magno-cellular pathway. This candidate information is in turn transmitted to the left ITS (inferior temporal sulcus) and LOS (lateral occipital cortex) as a top-down signal, and then the object is identified. Processes of mental rotation and object recognition were also examined psychologically and physiologically in details and the responsible regions in the cortex were identified. However it remains unclear how these functions are realized and integrated in the brain at the level of the neural network.

2 Structure of the Model

The parietal network is a bi-directional network consisting of three layers, the CIP layer, the AIP (anterior intra-parietal area) layer and the PM (premotor cortex) layer. The temporal network is an RBF (Radial Basis Function) network [6] (see Fig. 1). Binocular image information of a stick pattern from the visual cortex is input into these two networks. Once this signal from the visual cortex is received, it is matched with a shape in memory (target object), and then the angular difference φ of rotation from the target object is estimated by the RBF network of the temporal cortex. When the degree of matching is high enough, it is instantly identified. However, when the degree of matching is low, the angular difference φ from the target object is estimated, which is then transmitted to the PM layer. This information triggers the rotation command, which is transmitted from PM layer as a top-down signal, down to the visual cortex via CIP, and rotated image representation is internally generated in the visual cortex. Then this information of rotated image is transmitted again to RBF network. In this way, the rotated image generated by the top-down signal is checked by the RBF network for a match between the rotated shape and the target object. As the image in the visual cortex is rotated gradually over a number of time

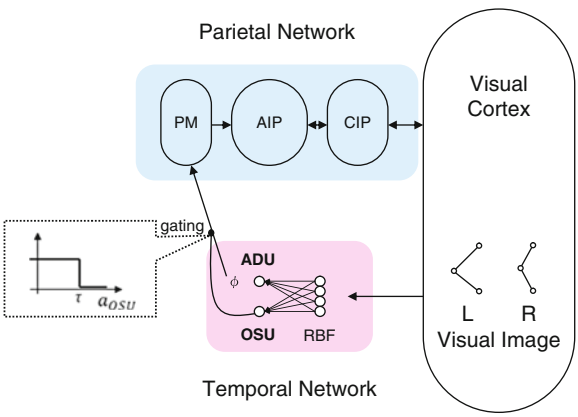


Fig. 1 Proposed scheme of the network model

steps, the matching level rises. When it is above threshold τ, the gating network comes into operation, and the mental rotation stops because of inhibition of signals from the temporal cortex (the output of the RBF network) to the PM layer.

3 Training Procedure of the Network

The parietal network is an extension of that presented by Ashizawa and Inui [7, 8]. This network was trained as follows: at time *t*, the parietal network is given both a stick pattern viewed from a certain direction to the input layer (visual cortex) and a rotation command to PM layer. Then the network is trained to generate the image rotated by a certain angle in the visual cortex at time $t + 1$. The angle is specified by the rotation command.

Images of the target object, viewed from various directions, are then presented to the temporal network (RBF network) and it is trained using a back-propagation learning algorithm. The output becomes 1.0 if we present the image of the target object; otherwise, it is 0.0. In this simulation, the number of images of the target object from various viewpoints was limited. As a result, the temporal network learns to output an intermediate value between 0 and 1 for any novel view of the target object.

4 Results and Discussion

First, we examined whether the parietal network can reproduce a precise mental rotation for a single bar; that is, whether the bar can be mentally rotated based on the rotation command from the PM. The results are shown in Fig. 2. Continuous curves in Fig. 2a represent the physical rotation in tilt-slant space for 8 different slopes in 3D space. Figure 2b–d show the mental rotation in 3D space of the bar (cross) for three different angular velocities **ω** with physical rotation (continuous curve).

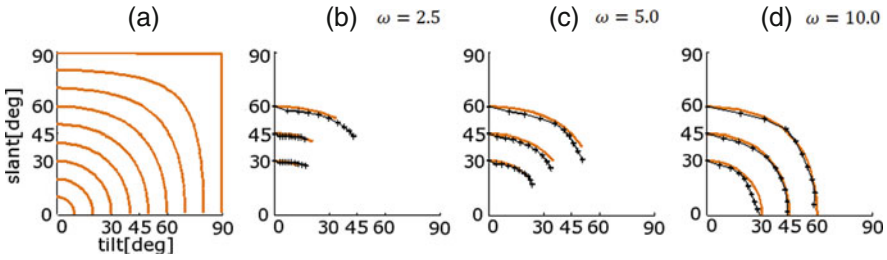


Fig. 2 Physical and reproduced rotation for different slopes of a bar in tilt-slant space. **ω** is the angular velocity

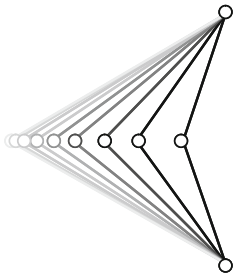


Fig. 3 Schematic response for the left monocular image in the visual cortex during mental rotation

Figure 2 indicates that the network correctly reproduces mental rotation of the bar in accordance with the rotation command. A schematic response of the visual cortex during mental rotation is also shown in Fig. 3.

After this learning, neurons in the CIP layer of the parietal network exhibit localized response selectivity in tilt-slant space that is similar to the physiological characteristics of the real CIP neurons [9]. In our model, movement of the activation pattern in tilt-slant space caused by the rotation command of PM layer corresponds to mental rotation.

In addition, the output of the neurons of the AIP layer is modulated by a motor command from PM layer, which is similar to the properties of neurons in the monkey parietal cortex. Furthermore, the model predicts that activation of the CIP is dependent on the rotation angle, which was also reported in a functional imaging study of mental rotation [3].

Acknowledgments This study was supported by Grant-in Aid for Scientific Research (S) "Brain mechanisms of dynamic image generation based on body schema" (20220003) from the Japanese Ministry of Education, Science, Technology, Sports and Culture (MEXT).

References

1. Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Chabris, C.F., Rauch, S.L., Anderson, A.K.: Identifying objects seen from different viewpoints: A PET investigation. Brain. **117** (1994) 1055–1071.
2. Sugio, T., Inui, T., Matsuo, K., Matsuzawa, M., Glover, G.H., Nakai, T.: The role of the posterior parietal cortex in human object recognition: a functional magnetic resonance imaging study. Neurosci. Lett. **276** (1999) 45–48.
3. Schenden, H.E., Stern, C.E.: Mental rotation and object categorization share a common network of prefrontal and dorsal and ventral regions of posterior cortex. NeuroImage. **35** (2007) 1264–1277.
4. Tarr, M.J.: Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. Psychon. Bull. Rev. **2** (1995) 55–82.
5. Kveraga, K., Boshyan, J., Bar, M.: Magnocellular projections as the trigger of top-down facilitation in recognition. J. Neurosci. **27** (2007) 13232–13240.
6. Poggio, T., Edelman, S.: A network that learns to recognize three-dimensional objects. Nature. **343** (1990) 263–266.
7. Ashizawa, M., Inui, T.: Neural network models of mental rotation. The 6th Conference of the

Japanese Society for Cognitive Psychology (2008).
8. Ashizawa, M., Inui, T.: A bi-directional neural network model of 3-D mental rotation. IEICE Tech. Rep. **TL2008-45** (2008) 7–12.
9. Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., and Tsutsui, K.: Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. Philosophical Transactions of the Royal Society of London. Ser. B Biol. Sci. **353** (1998) 1363–1373.

Representation of Time-Series by a Self-Similar Set in a Model of Hippocampal CA1

Yutaka Yamaguti, Shigeru Kuroda, and Ichiro Tsuda

Abstract Because episodic memory includes a time series of events, an underlying dynamics for the formation of episodic memory is considered to employ a mechanism of encoding sequences of events. The "Cantor coding" hypothesis in hippocampal CA1 has been proposed, which provides a scheme for encoding temporal sequences of events. Here, in order for investigating the Cantor coding in detail, we constructed a model for the CA1 network which consists of conductance-based model neurons. It was assumed that the CA3 outputs temporal sequences of spatial patterns to CA1. It was shown that the output patterns of CA1 were hierarchically clustered in a self-similar manner according to the similarity of input time series. The dependency of the efficacy of encoding on the input time interval and its robustness against noise was investigated.

Keywords Hippocampus · Cantor coding · Iterated function system

1 Introduction

A consensus has emerged that the hippocampal formation is crucial for episodic memory. For episodic memory, representing a temporal sequence of events is one of important factors. The hippocampal CA3 region has dense recurrent connections between pyramidal neurons. Since the seminal work of D. Marr [1], the CA3 has been considered to be an associative memory [2, 3]. Recently, several experimental supports for this idea appeared [4, 5]. On the other hand, some researchers suggest that the CA1 region is tuned for the processing of temporal information [6].

Here we investigate the information representation in CA1 neurons when the CA1 network receives time-dependent activity from CA3. Various kinds of rhythmic

Y. Yamaguti (✉)
Research Institute for Electronic Science, Hokkaido University, Sapporo 001-0020, Hokkaido, Japan; Research Center for Integrative Mathematics, Hokkaido University, Sapporo 001-0020, Hokkaido, Japan
e-mail: yamaguchi@math.sci.hokudai.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_14,

activity have been observed in the hippocampus and its neighborhood [7, 8]. Especially, theta and gamma oscillations have been observed in episodic context. Thereby, continuously input stimuli may synchronize with these rhythms. From these considerations, in this study, we assume that the input time-series from CA3 to CA1 is discrete, but not continuous one. In other words, we assume that the input time-series consists of several elementary spatial patterns of activities of neurons. The hypothesis of "Cantor coding" in CA1 has been proposed by Tsuda and Kuroda [9]. Cantor coding provides the scheme for encoding temporal sequences of events. It forms a fractal-like hierarchical structure in the state space of neural dynamics. Tsuda and Kuroda propose a simple CA3–CA1 model in which CA1 encodes sequences. In physiological experiments by Fukushima et al. [10], hierarchical representations of input time-serieses in the membrane potential of CA1 neuron were actually observed. This finding suggests the presence of Cantor coding in the actual hippocampus. Therefore, here we try to develop a biology-oriented model of CA1 in order to study the possibility of Cantor coding in CA1. Then we investigate the physiologically-plausible condition in which Cantor coding can emerge and provide some predictions on possibility of observation of Cantor coding in experimental studies.

2 Model

We propose a model of the CA1 network for Cantor coding. Figure 1 shows a schematic drawing of our CA1 network receiving sequential inputs from CA3. As a first step, in order to clarify the role of excitatory pyramidal neurons, here we neglect interneurons. Because the fact that there are much less recurrent synaptic connections in CA1 compared with CA3, we neglect such recurrent connections. The CA3 network is supposed to store M spatial firing patterns $\{X^0, X^1, ..., X^{M-1}\}$, which we call elementary patterns. The model CA1 consists of N pyramidal neurons. These neurons receive temporal series of spatial patterns via AMPA- and NMDA-receptor mediated synaptic activity. We consider the situation that rhythmically switching sequential pattern is delivered to the CA1 network. One of $m(< M)$ elementary input patterns defined above is randomly chosen every T-ms (Fig. 1b). We used two-compartment model proposed by Pinsky and Rinzel [11] for a model of a pyramidal neuron in the hippocampus.

3 Cantor Coding

We investigated the dynamic behavior of membrane potentials and output spike patterns produced by the model CA1 neurons. In both distribution of membrane potentials and spike patterns, we found Cantor sets. In order to understand how the distribution of output spike patterns is formed in the high-dimensional space, we utilize principal component analysis (PCA), transforming high-dimensional vectors

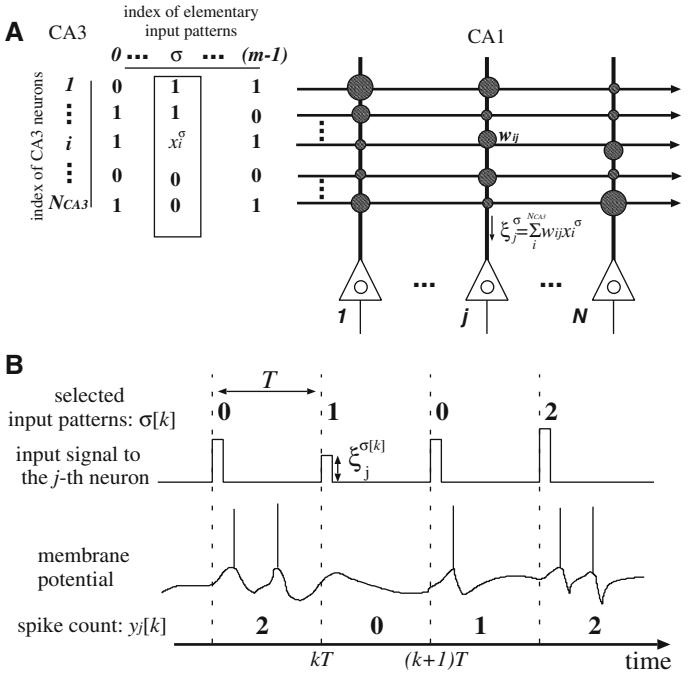


Fig. 1 Schematic illustration of the CA1 network. **a** inputs from CA3, and the CA1 network receiving such inputs. **b** representation of input time series to the *j*-th CA1 neuron and response of the CA1 neuron

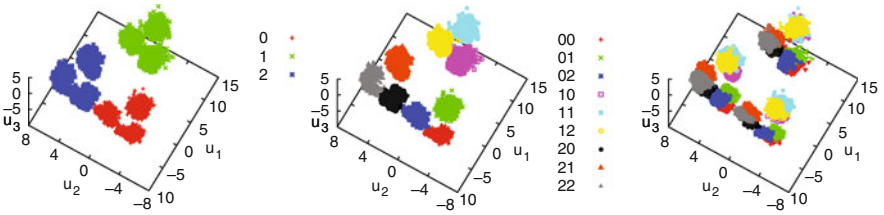


Fig. 2 Scatter plots of the output vectors of the CA1 neurons. Each *symbol* indicates an spatial firing pattern just after receiving an particular input sequences of length sequences one (*left*), two (*middle*) and three (*right*) denoted by legends. An input time interval $T = 100$ and a number of input elementary pattern $m = 3$

to a lower-dimensional space. Figure 2 is a typical example of self-similar structure emerging in a space of output vectors.

Furthermore, using linear discriminant analysis, we clarified that the network shows good performance for discriminating input sequences with the length of at least five, and that the performance of coding sensitively depends on the input strength and the interval of input sequence. Coding efficacy is quite good when the interval T is 80–200 ms, which is a characteristic time window of this system

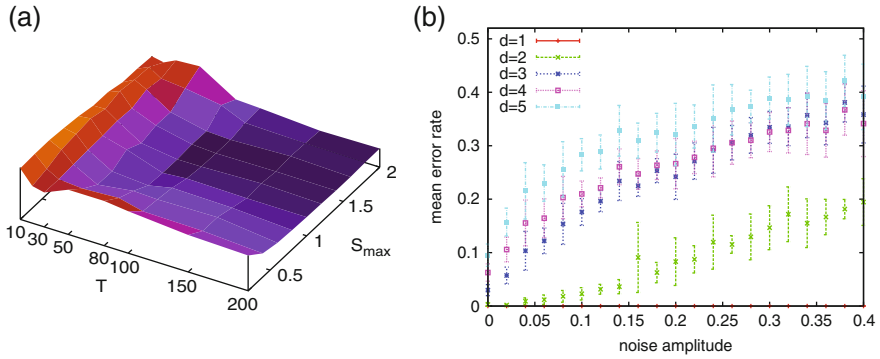


Fig. 3 **a** Efficacy of the coding as a function of input time interval T and input strength $S_{\max}$. z-value is a ratio of misclassification of a classifier constructed by linear discriminant analysis. Lower values corresponds to better encoding efficiency. Length of input sequences to to be classified, $d = 3$. **b** Coding efficacy as a function of strength of noise

(Fig. 3a). These results indicate the possibility of decoding from a pulse train output of CA1 pyramidal neurons. The encoding scheme has a certain degree of tolerance against noise (Fig. 3b).

4 Summary

In this study, we clarified the presence of Cantor coding for the randomly input spatio-temporal patterns in the CA1 network consisting of neurons described by the two-compartment model. The distribution of spatial firing patterns shows fractal-like clustering with hierarchical structure, which reflects the similarity of input time series. The coding performance and its dependence on control parameters were systematically investigated. We found time windows for Cantor coding.

Acknowledgments This work was supported in part by Grants-in-Aid (No. 18019002 and 18047001) for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the twenty-first century COE Program of the Department of Mathematics of Hokkaido University.

References

1. Marr, D.: Simple memory: A theory for archicortex. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **262**(841) (1971) 23–81.
2. Mcnaughton, B.L., Morris, R.G.M.: Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends. Neurosci. **10**(10) (1987) 408–415.
3. Treves, A., Skaggs, W.E., Barnes, C.A.: How much of the hippocampus can be explained by functional constraints? Hippocampus **6**(6) (1996) 666–674.
4. Nakazawa, K., Quirk, M.C., Chitwood, R.A., Watanabe, M., Yeckel, M.F., Sun, L.D., Kato, A., Carr, C.A., Johnston, D., Wilson, M.A., Tonegawa, S.: Requirement for hippocampal CA3 NMDA receptors in associative memory recall. Science **297**(5579) (2002) 211–218.

5. Wills, T.J., Lever, C., Cacucci, F., Burgess, N., O'Keefe, J.: Attractor dynamics in the hippocampal representation of the local environment. Science **308**(5723) (May 2005) 873–876.
6. Kesner, R.P., Lee, I., Gilbert, P.: A behavioral assessment of hippocampal function based on a subregional analysis. Rev. Neurosci. **15**(5) (2004) 333–351.
7. Buzsáki, G.: Theta oscillations in the hippocampus. Neuron **33**(3) (2002) 325–340.
8. Montgomery, S.M., Buzsáki, G.: Gamma oscillations dynamically couple hippocampal CA3 and CA1 regions during memory task performance. PNAS **104**(36) (2007) 14495–14500.
9. Tsuda, I., Kuroda, S.: Cantor coding in the hippocampus. Japan. J. Ind. Appl. Math. **18**(2) (2001) 249–258.
10. Fukushima, Y., Tsukada, M., Tsuda, I., Yamaguti, Y., Kuroda, S.: Spatial clustering property and its self-similarity in membrane potentials of hippocampal CA1 pyramidal neurons for a spatio-temporal input sequence. Cogn. Neurodyn. **1**(4) (2007) 305–316.
11. Pinsky, P.F., Rinzel, J.: Intrinsic and network rhythmogenesis in a reduced traub model for CA3 neurons. J. Comp. Neurosci. **1**(1–2) (1994) 39–60.

Emergence of Iterated Function Systems in the Hippocampal CA1

Shigeru Kuroda, Yasuhiro Fukushima, Yutaka Yamaguti, Minoru Tsukada, and Ichiro Tsuda

Abstract In rat CA1 pyramidal cells, we previously observed hierarchical clusters of the distribution of membrane potentials, arranged according to the history of input sequences. In this study, we deal with the dynamical mechanism generating such a hierarchical distribution. The recording data were investigated using return map analysis. Each of the obtained return maps was well approximated by a set of contractive affine transformations. These findings provide direct evidence that the information of temporal sequences generated in CA3 can be self-similarly represented in the membrane potentials of CA1 pyramidal cells.

1 Introduction

By the clinical studies, it is established that the hippocampus is a necessary organ of the formation of episodic and semantic memories, especially for episodic memory. The hippocampus receives all kinds of sensory information via entorhinal cortex. One of the main components of hippocampus, CA3, is considered to function as a network for autoassociative memories via the framework of attractor dynamics, where memories can be stably stored as corresponding neuronal patterns and can be retrieved by partial cues. Since pyramidal cells in CA1 area have less recurrent connections compared with CA3, it may be thought that CA1 has different functional roles form CA3. One hypothesis is that CA1 would be involved in the information processing of spatiotemporal sequence from CA3. We have propose a scheme for encoding the temporal sequence of events in CA1, which we refer to as "Cantor coding" [1] and have discussed its significance for the formation of episodic memory in the hippocampus-cortex system [2, 3]. Cantor coding enables the temporal pattern

S. Kuroda (✉)
Research Institute for Electronic Science, Hokkaido University, Sapporo 001-0020, Hokkaido, Japan
e-mail: kuroda@math.sci.hokudai.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_15,

sequences generated in CA3 to be represented hierarchically in fractal subsets in state space of CA1. How can we verify the hypothesis of Cantor coding in CA1?

We conducted experiments using rat hippocampal slices to clarify how the spatiotemporal sequence delivered via Schaffer collaterals affects the postsynaptic membrane potentials of individual hippocampal CA1 pyramidal cells. We found that the membrane potentials of CA1 pyramidal cells were hierarchically clustered according to the histories of input sequences up to two or three length [4]. However, the finding of such hierarchical clusters is still only indirect evidence of the presence of Cantor sets, because these sets are essentially infinite objects; observed sets are finite. A direct evidence of Cantor sets may be obtained by showing the existence of emergent rules such as iterated function systems (IFSs) which provide a deterministic framework for generating self-similar fractal patterns as their attractors. From this point of view, the experimental data was investigated using return map analysis [5]. We also deal with a collective behavior at population level, using a reconstructed multi-cell recording data set.

2 Materials and Methods

Patch-clamp recordings were made from pyramidal cells of CA1 area in rat hippocampal slices (Fig. 1a). In order to generate spatiotemporal inputs to the pyramidal cell in CA1, two stimulus electrodes were set to the Schaffer collaterals, in sites proximal and distal to the soma, respectively. For each cell, a recording session consisted of 122 stimulus periods with an intervening rest period of 10 s. In a stimulus period, ten successive inputs were applied with 30 ms intervals. Each input pattern was randomly selected among the four spatial input patterns of electrical stimulations: both electrodes ("4"), a electrode placed in the proximal site ("3"), or the distal site ("2"), and neither electrode placed ("1").

We recorded from 11 cells in 6 slices. One cell was excluded from analysis because the cell exceeded a criterion for variation of resting potential during the recording session. The ten cells were classified into two groups, *sub-threshold* (cell1, ..., cell5) and *supra-threshold* (cell6, . . ., cell10), according to whether or not the continual stimulations induced spikes. For each stimulus-period, the baseline membrane potential was determined as mean amplitude during 2 s before the stimulus period. Hereafter, we express membrane potential as the difference between the measured voltage and the baseline membrane potential at each stimulus period.

A response at Δt to nth input was defined as the membrane potential at a fixed elapsed time Δt after the input, which is denoted by $V_{\Delta t}(n)$ (Fig. 1b). In particular, $V_{\mathrm{last}}(n)$ denotes the value at $\Delta t = 28$ ms taken as the timing just before the next input. Responses for analysis were gathered from all stimulus periods for each cell using the same procedure. Return map analysis was used to examine the dynamics underlying the generation of responses to a spatiotemporal input sequence. For a response sequence $\{V_{\mathrm{last}}(n)\}_n$, a return map was generated by plotting each response $V_{\mathrm{last}}(n)$ against the previous response, $V_{\mathrm{last}}(n-1)$.

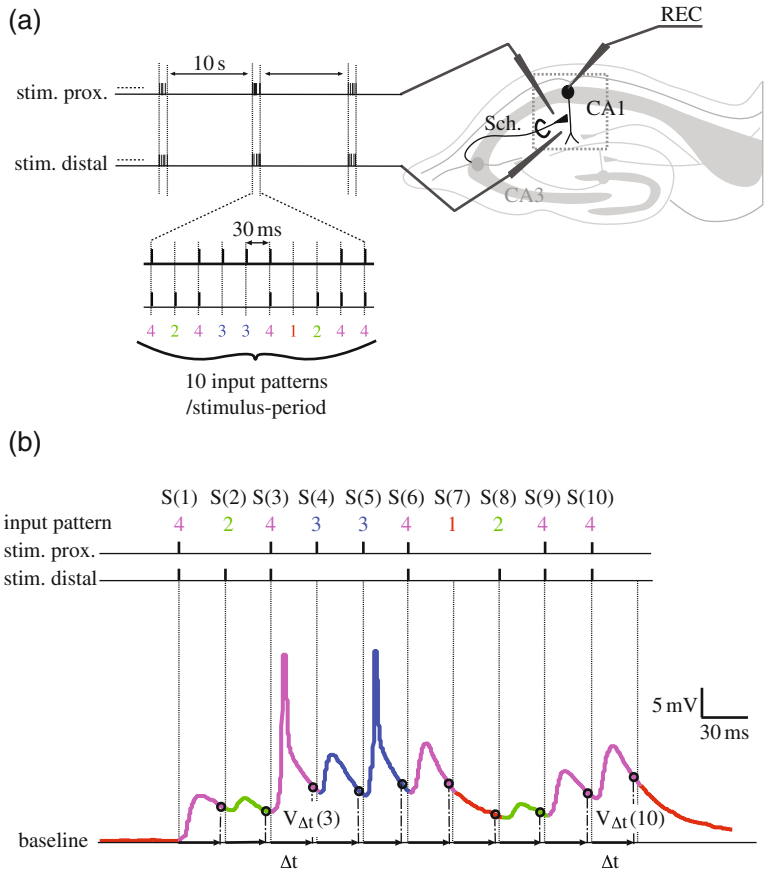


Fig. 1 Experimental procedure and sample traces. **a** Schematic diagram of the experimental procedure. In each stimulus period, input patterns were randomly selected among four spatial input patterns of electrical stimulations: "1" (*red*), "2" (*green*), "3" (*blue*) and "4" (*magenta*). In the figures in this article, these input patterns are *color-coded* in the same way throughout. **b** Sample traces of membrane potentials and the neuronal responses in a stimulus period. The *upper two samples* show the timing of electrical stimulation and the *lower sample* shows the timing of responsive membrane potentials recorded from a cell in supra-threshold group, smoothed using a median filter to remove electrical stimulus artifacts. The membrane potentials are *color-coded* according to the kinds of their most recent input patterns (Colors will be viewed in online version)

3 Results and Conclusions

At individual cell and population levels, a return map of the response sequence of CA1 pyramidal cells was well approximated by a set of contractive affine transformations (Fig. 2). These finding strongly suggest that CA1 dynamics receiving spatiotemporal input from CA3 has a mode that is characterized by input-driven IFS consisting of a self-organized response rule for each spatial input pattern. This

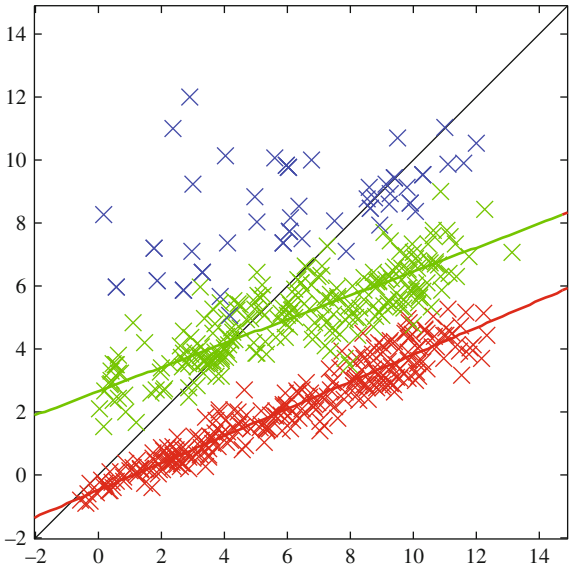


Fig. 2 Return map of response sequence V_{last} at individual cell level. An example of a cell in supra-threshold group. $V_{\text{last}} := V_{\Delta t} = 28$[ms] The return map consists of four parts, called *branches*, corresponding to four $n + 1$st input patterns. The *colors of the points* indicate the kinds of $n + 1$st input patterns. The points $V_{\text{last}}(n)$, $V_{\text{last}}(n + 1)$ such that spikes occur in the $n + 1$st input interval are enclosed by *open circles*. Superimposed on the branches are the fitting lines using major axis regression and the diagonal line $V_{\text{last}}(n + 1) = V_{\text{last}}(n)$. Successive responses, $V_{\text{last}}(n)$ and $V_{\text{last}}(n+1)$, in each branch had a decent correlation coefficient, and all slopes of fitting lines for the branches were smaller than 1. These indicate the presence of contractive affine transformations (Colors will be viewed in online version)

dynamics ensures that the distribution of response is hierarchically clustered according to input histories, and also ensures that a spatial and retrospective code table can be automatically formed. Hence we obtain Cantor coding.

References

1. Tsuda, I., Kuroda, S.: Cantor coding in the hippocampus. Jpn. J. Ind. Appl. Math. **18** (2001) 249–281.
2. Tsuda, I.: Towards an interpretation of dynamic neural activity in term of chaotic dynamical systems. Behav. Brain Sci. **24**(5) (2001) 793–847.
3. Tsuda, I., Kuroda, S.: A complex systems approach to an interpretation of dynamic brain activity II: does cantor coding provide a dynamic model for the formation of episodic memory. Lect. Notes Comput. Sci. **3146** (2004) 129–139.
4. Fukushima, Y., Tsukada, M., Tsuda, I., Yamaguti, Y., Kuroda, S.: Spatial clustering property and its self-similarity in membrane potentials of hippocampal CA1 pyramidal neurons for a spatio-temporal input sequence. Cogn. Neurodyn. **1**(4) (2007) 305–316.
5. Kuroda, S., Fukushima, Y., Yamaguti, Y., Tsukada, M., Tsuda, I.: Iterated function systems in the hippocampal CA1. Cogn. Neurodyn. **3**(3) (2009) 205–222.

Frontal Theta for Executive Functions and Parietal Alpha for Storage Buffers in Visual Working Memory

Masahiro Kawasaki and Yoko Yamaguchi

Abstract Theta and alpha brain rhythms play an important role in visual working memory (VWM), while their individual functions are not certain. This study conducted the EEG experiments for a mental manipulation task. The frontal theta and parietal alpha activity increased for manipulation and maintenance of visual information in minds, respectively. Furthermore, phase synchronization between the frontal and parietal areas was significantly increased. These results suggested that the communication between the frontal theta activity for central executive and the parietal alpha activity for storage buffers would mediate the VWM.

Keywords Working memory · Frontal · Parietal · Theta · Alpha · Synchronization

1 Introduction

In everyday life, we can temporarily retain visual information from the external world and manipulate the mental representation in our mind. Such a high-level brain function is executed by visual working memory (VWM), which includes the aspect of not only passive short-term maintenance ("storage buffer") but also active manipulation ("central executive") such as control of the allocation of attention, selection of relevant information, and transformation of the mental representations.

To investigate the neural substrate for VWM, previous electrophysiological studies on primates and neuroimaging studies on humans have shown that a vast network of brain regions including frontal, parietal, and visual cortices exhibit sustained activity during the retention interval of delayed-matching tasks [1]. Especially, the posterior parietal cortex is proposed to play a general role in maintaining the actual contents of representations, since the delay-period activity correlated with the

M. Kawasaki (✉)
Rhythm-based Brain Computation Unit, RIKEN BSI-TOYOTA Collaboration Center, Saitama, Japan; Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute, Saitama, Japan
e-mail: kawasaki@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_16,

number of various items held in VWM (i.e., VWM load) [2]. On the other hand, the prefrontal cortex is thought to act as an attentional filter – determining which information within the visual scene will be maintained and updating them in VWM, and not participating in the actual maintenance of that information [1]. Indeed, the frontal cortex shows the transient activity during selection period and not any sustained activity during the retention interval [3], and further the frontal delayed activity is sensitive to VWM load only when selection of relevant information is required [4].

Thus, although there are rich neurological evidences for VWM, the important question remains as to how the central executive accesses the storage buffer for manipulation of VWM. To approach the issue, the temporal relationship of electroencephalograph (EEG) oscillations between the specific brain areas is of particular useful. Previous EEG studies have shown the modulated theta and alpha rhythms during VWM tasks [5], but have not yet clearly identified these functional roles and relationships. Here, we investigated the dynamic links between the central executive and storage buffers, by using time-frequency analyses of EEG data recorded during a mental manipulation task.

2 Material and Methods

Fourteen healthy volunteers (4 females; mean age = 27.9 ± 6.8 years) took part in the EEG experiment, after providing informed consent. At the beginning of each trial, 5 × 5 gridded squares and a red circle included within one of those squares presented on the computer screen for 1 s (sample display, Fig. 1). In one trial, participants were first required to memorize, and then maintained the circle position in the gridded squares. After that, a white arrow designating a direction (upward, downward, rightward, or leftward) to which the participants should move the red circle by one square in their minds was presented for 1 s. Participants must repeat the arrow based mental manipulation for four times, and then determine via button press whether the position of the circle which they mentally moved matched a probe visual stimulus (test display). The duration of the inter-trial interval (ITI) was 2 s. All participants completed 24 trials. They trained the task before the EEG sessions.

EEG was recorded using the 64 scalp electrodes (sampling rate: 500 Hz), amplified by using the NeuroScan equipment, and filtered in the band-pass range from 0.1 to 50 Hz. Reference electrodes were placed on the right and left ears. We segmented EEG data for the correct trials to 3 s-epochs for manipulation period from the onset of the instruction. The independent components analyses components which

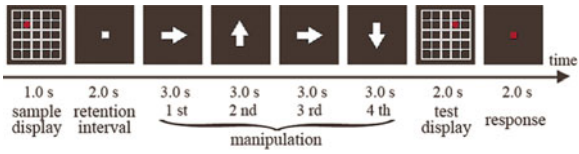


Fig. 1 Schematic illustrations of one trial sequence for the mental manipulation task

were significantly correlated with the EOG were eliminated from the data as the artifact involved in eye blinks and movements. And further, to reduce the effect of the spreading cortical currents due to volume conductance producing voltage maps covering large scalp areas, current source density conversion was performed.

Time-frequency (TF) amplitudes and phases were calculated by wavelet transforms implementing the convolution of the original EEG signals with complex Morlet's wavelet function of 7-cycles length, with frequency ranging from 1 to 20 Hz (0.5-Hz steps). The TF amplitudes for manipulation were corrected by subtracting the ITI amplitudes and averaged across trials.

To identify the phase relations between any two electrodes, the phase synchronization index (PSI) for each electrode-pair was defined by the equation:

$$\mathrm{PSI}_{jk}(t,f) = \sqrt{\left(\sum_{i=1}^{N}\cos(\Delta\phi_{jk}(i,f)/N)\right)^2 + \left(\sum_{i=1}^{N}\sin(\Delta\phi_{jk}(i,f)/N)\right)^2} \quad (1)$$

where $\Delta\Phi_{jk}(t,f)$ is the phase differences between jth and kth electrodes, and N is number of time points. We applied the bootstrap to the PSIs of the individual subjects and compared the virtual PSI data during the manipulation and ITI periods [6].

3 Results

All participants achieved high accuracy rate for the tasks (mean rate = 97.1 ± 4.7%). The averaged TF amplitudes for the manipulation periods relative to those of ITI showed that the theta (4–6 Hz) amplitudes increased in the frontal and parietal areas whereas the alpha ones (10–12 Hz) were enhanced in only the parietal areas (Fig. 2a, b). Furthermore, according to the duration period of theta and alpha in manipulation period, we could find that early active manipulation is followed by a maintenance period. The increased parietal theta and alpha activity was also survived but the frontal theta activity disappeared during the maintenance period (data not shown). The theta (6 Hz) and alpha (12 Hz) PSIs between the frontal and parietal electrodes were significant ($Z = 3.70$, $p < 0.01$), indicating that phase synchronization during manipulation period was higher than the ITI.

4 Discussion

Our results of the TF amplitudes demonstrated a clear dissociation within a distributed VWM network between the functions; the central executive and storage buffers. The frontal theta activity was observed in the manipulation not in the maintenance period. On the other hand, the parietal alpha activity was enhanced

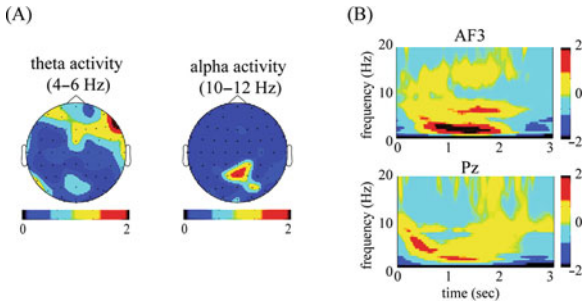


Fig. 2 EEG results. **a** Topographical maps of theta and alpha amplitudes during manipulation. **b** Time-frequency graphs of the frontal (AF3) and parietal (Pz) electrodes

in both the manipulation and maintenance periods. These results support the idea that the frontal cortex is associated with active manipulation and the parietal cortex is involved in maintenance [1–4]. In addition, these results also clarified different roles of different local synchronization within the specific cell assemblies, in VWM function. The frontal and parietal theta can predict active manipulation possibly including interplay between the two regions. The parietal alpha can predict maintenance that continues in the manipulation period.

Both the theta and alpha phase synchronizations between the frontal and parietal areas during manipulation periods were higher than the ITI. A few studies have independently reported the increased theta and alpha phase synchronization during working memory tasks [5, 6]. In this study, the co-occurrence of theta and alpha phase synchronization, which reflects co-activation of two functional systems, provides the neurological evidence that the central executive would access the storage buffers by the global synchronization.

In summary, our findings about oscillatory amplitude of the theta and alpha rhythms and the phase synchronization between the two distinct areas have clearly shown that the dynamic linking between the theta and alpha oscillations would mediate communication between the central executive and storage buffers in VWM. Future studies should clarify a cross-frequency coupling between the theta phases and alpha phases among the frontal and parietal areas.

References

1. Curtis, C.E., D'Esposito, M.: Persistent activity in the prefrontal cortex during working memory. Trends Cogn. Sci. **7** (2003) 415–423.
2. Todd, JJ., Marois, R.: Capacity limit of visual short-term memory in human posterior parietal cortex. Nature. **428** (2004) 751–754.
3. Rowe, J.B., Toni, I., Josephs, O., Frackowiak, R.S.J., Passingham, R.E.: The prefrontal cortex: response selection or maintenance within working memory?. Science. **288** (2000) 1656–1660.
4. Kawasaki, M., Watanabe, M., Okuda, J., Sakagami, M.: SFS for feature selective maintenance, IPS for simple maintenance in visual working memory. J. Vis. **6** (2008) 1093.
5. Klimesch, W., Freunberger, R., Sauseng, P., Gruber, W.: A short review of slow phase synchronization for control processes in different memory systems?. Brain Res. **1235** (2008) 31–44.
6. Mizuhara, H., Wang, L.Q., Kobayashi, K., Yamaguchi, Y.: Long-range EEG phase-synchronization during an arithmetic task indexes a coherent cortical network simultaneously measured by fMRI. NeuroImage. **27** (2005) 553–563.

Associative Latching Dynamics vs. Syntax

Eleonora Russo, Sahar Pirmoradian, and Alessandro Treves

Abstract We model the cortical dynamics underlying a free association between two memories. Computationally, this process may be realized as the spontaneous retrieval of a second memory after the recall of the first one by an external cue, what we call a *latching transition*. As a global cortical model, we study an associative memory Potts network with adaptive threshold, showing latching transitions. With many correlated stored patterns this unstable dynamics can proceed indefinitely, producing a sequence of spontaneously retrieved patterns. This paper describes the informational properties of latching sequences expressed by the Potts network, and compares them with those of the sentences comprising the corpus of a simple artificial language we are developing, BLISS. Potts network dynamics, unlike BLISS sentences, appear to have the memory properties of a second-order Markov chain.

Keywords Memory · Attractor dynamics · Information · Artificial language

1 Introduction

Cortical networks have been thought to retrieve memories associatively [1], both at the local and global level [2]. The simple Hopfield neural network model [3] has stimulated the study of content-addressed, auto-associative retrieval in terms of attractor dynamics. Once a memory has been retrieved by an external cue, however, if the corresponding attractor state is made unstable, it may serve itself as an internal cue for a second memory. Based on this simple observation, many authors have explored recurrent networks that model processes of the free association between two memories. These proposals differ in the ingredients introduced in order to destabilize the first attractor state, so as to produce a spontaneous sequential retrieval of several stored memories.

E. Russo (✉)
SISSA Cognitive Neuroscience, Trieste, Italy
e-mail: russo@sissa.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_17,

In 1986 Sompolinsky and Kanter [4] proposed a simple network that, with a specific set of connection weights, could retrieve an equally specific sequence of memories. In 1987 Tsuda [5, 6] proposed a model with two coupled networks, a stable and an unstable one, interacting with each other and generating oscillating retrieval. Hermann et al. [7], in 1993, obtained "episodic" and "semantic" transitions among memories with a dynamic threshold influencing active neurons.

Similarly to the work of Hermann, we study here a network where transitions among patterns are due to an adaptive threshold, but with a different kind of units.

2 Potts Model

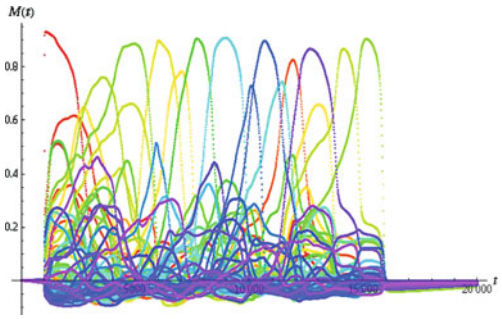
The network, described in detail in [8], is a recurrent network with Potts units. Instead of a single neuron each unit represents a local cortical network, and can take S different uncorrelated active states (plus one inactive state). The active states represent memories at the local network level, realized as local attractor states through a process the model does not describe, as it focuses on the global cortical level. Within each active state the activation rate is taken to be threshold-linear. We study the global attractor dynamics, after storing on the tensor connections among Potts units p global activity patterns. Each unit receives C connections.

To have spontaneous transitions among attractors we introduce a fatigue model: both an overall and a state-specific adaptive threshold which, with a characteristic time constant, tracks the mean activation value of each Potts unit in that state. After retrieving a global pattern the system is in a quasi-stable attractor, as the threshold continues to increase and to weaken each active state, until the unit changes state or becomes inactive. The free-energy landscape of configurations is then dynamically modified, leading the system to abruptly jump to a new attractor, often nearby.

In order to have structured latching transitions, however, the introduction of a dynamic threshold is not enough: correlations among patterns are also needed. In [9] we have studied the detailed dynamics of transitions between two correlated patterns. In this simplified condition we have identified three types of latching transition, each characterized by a range of correlations: quasi-random transitions between weakly correlated attractors, history-dependent transitions between attractors with stronger correlations, oscillatory transitions between pairs of closely overlapping attractors.

3 Extended Potts Dynamics

With a generative algorithm, as explained in [8], we produce a set of correlated patterns and we store them into the network. In this more natural and complex condition, transitions among all pairs of patterns may occur, and they still seem to cluster in the three types above. Starting with an external cue, that induces the retrieval of



→ Ahriman takes Zarathustra

→ New lions build the house

→ The cow dances

→ The knife assumes that the worm sees AhuraMazda

→ Yasmin knows that Magi don't believe that the ox doesn't need the dog

→ Great black stones give the stone to Yasmin

Fig. 1 (*left*) Example of Potts dynamics: different *colors* show the overlap of the network with different patterns; (*right*) Examples of sentences generated by the artificial language BLISS (Colors will be viewed in online version)

a pattern, a series of transitions follows, as in Fig. 1, until activity dies out or we stop the simulation. Its duration increases with the correlation among patterns and with p, the total number of patterns. Latching dynamics can be quite disorderly, but we can extract the sequence of patterns that, at each time t, have the highest overlap with the activity of the Potts network. We can then study the properties of such discrete sequences, neglecting the originally continuous nature of the dynamics.

4 BLISS Sentences

In a separate project, we have designed an artificial Basic Language Incorporating Syntax and Semantics, BLISS, in order to test the language acquisition capability of the Potts and of other networks. BLISS is intended to be of intermediate complexity, and in its current provisional form it includes a stochastic regular grammar with 30 production rules but no semantics yet. The associated probabilities (e.g. for transitive vs. intransitive verbs) are fine-tuned to the statistics of the one-million-word Wall Street Journal (WSJ) corpus. The 170 terminal symbols belong to different lexical categories such as verb, noun, adjective, determiner, demonstrative, . . ., whose relative frequencies are also tuned to the WSJ corpus. With a Perl code we have generated 50,000 BLISS sentences, of length between 5 and 12 words and with maximum 3 levels of embedding.

- Ahriman takes Zarathustra
- New lions build the house
- The cow dances
- The knife assumes that the worm sees AhuraMazda
- Yasmin knows that Magi don’t believe that the ox doesn’t need the dog
- Great black stones give the stone to Yasmin

5 Entropy and Information

We may now focus on the memory properties of both processes, Potts latching dynamics and BLISS sentences, described as discrete sequences of "symbols" picked from among $p = 80$ memory patterns and $p = 170$ words, respectively.

In the Potts case, we have run $p \times p$ latching sequences stopping each sequence after 30 transitions, of which we have analysed the last 20. Averaging across all transitions, we count the occurrence of all possible single patterns, consecutive pairs and triplets of patterns, and compute e.g. the entropy $H(x_n)$ and the information among the symbols (= patterns) in position n and the ones in position $n+i$, $I(x_n;\ x_{n+i})$. We proceed similarly with BLISS, using sentences with at least 7 symbols (= words).

Results show that both systems span an entropy lower than $\log_2 p$, naturally for BLISS and in the Potts case due to finite size effects. Both have favourite pairs and triplets: while many among all the possible combinations never occur, others are frequently present. Comparing the frequency of each pair with that of its inverse, BLISS is seen to be almost fully asymmetric (either one never occurs), while Potts dynamics are substantially symmetric, as expected. The most frequent Potts pairs are comprised of strongly correlated patterns. For BLISS, in Fig. 2 (right) each color denotes the entropy at a specific position and the information which that position conveys to the next positions. Over half of the total variability at each position is seen to be independent of the words in previous positions, while the rest is determined by a long history of all preceding words, pointing at extended syntactic dependences.

In order to better characterize information propagation along the Potts latching sequence, in Fig. 2 (left) we have compared our sequences with Markov chains of first and second order. To equalize the noise present in a finite sample, we have also produced $p \times p$ Markov chains of 20 steps, with the two transition matrices of the

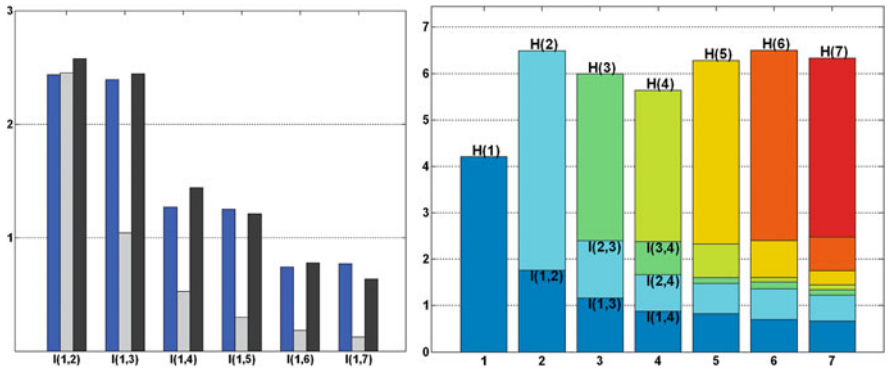


Fig. 2 (*left*) Information flow in the Potts latching sequence (*blue*), in a first-order Markov chain (*gray*) and in a second-order Markov chain (*black*); (*right*) Information flow in BLISS sentences, each *color* denoting the decaying influence of the word chosen at each position. Note the different *y*-scales (both in bits) (Colors will be viewed in online version)

first and second order chains extracted from the original Potts latching sequences. We can see in Fig. 2 that, whereas in a first order Markov process information decays faster than in BLISS, but also monotonically, the surprising trend of the latching information is similar to the one of a second-order Markov chain.

References

1. Marr, D.: A theory for cerebral neocortex. Proc. R. Soc. Lon B. **176** (1971) 161–234.
2. Braitenberg, V., Schüz, A.: Anatomy of the Cortex. Berlin: Springer-Verlag (1991).
3. Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U S A. **79** (1982) 2554–2558.
4. Sompolinsky, H., Kanter, I.: Temporal association in asymmetric neural networks. Phys. Rev. Lett. **57** (1986) 2861–2864.
5. Tsuda, I., Körner, E., Shimizu, H.: Memory dynamics in asynchronous neural networks. Prog. Theor. Phys. **78** (1987) 51–71.
6. Tsuda, I.: Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. Behav. Brain Sci. **24** (2001) 793–847.
7. Herrmann, M., Ruppin, E., Usher, M.: A neural model of the dynamic activation of memory Biol. Cybern. **68** (1993) 455–463.
8. Treves, A.: Frontal latching networks: a possible neural basis for infinite recursion. Cogn. Neuropsychol. **21** (2005) 276–291.
9. Russo, E., Namboodiri, V.M.K., Treves, A., Kropff, E.: Free association transitions in models of cortical latching dynamics. New J. Phys. **10** (2008) 015008.

Category Inference and Prefrontal Cortex

Xiaochuan Pan and Masamichi Sakagami

Abstract We hypothesize that the prefrontal cortex is involved in category inference. To test this hypothesis, we recorded single-unit activity from the lateral prefrontal cortex (LPFC) of two monkeys performing a sequential paired-association task with an asymmetric reward schedule. We found that a group of LPFC neurons encoded reward value specific to a category of visual stimuli defined by relevant behavioral responses. And these neurons predicted the amount of reward based on new category members that had never been used in the asymmetric reward task, when a member from the same category was paired with a large (or small) amount of reward. The results suggest that LPFC neurons encode category-based reward information, and transfer this information to category members including new ones, which could be the neural basis of category inference.

Keywords Category · Inference · Prefrontal cortex · Neuron · Monkey

1 Introduction

Category inference is a fundamental cognitive function that allows animals to relay information associated with one category member to others in the same group, even to novel ones solely on the basis of category membership. Behavioral studies have demonstrated that animals are capable of category inference [1, 2]. For instance, rats can transfer reward information associated with a particular stimulus-sequence (e.g. ABA, BAB) to new stimulus-sequences (e.g. CDC or DCD), because these sequences share the same temporal structure [2]. But little is known about neural mechanisms underlying category inference.

Several lines of evidence suggest the LPFC may be involved in category inference. First, the LPFC anatomically connects with the inferotemporal cortex, the

X. Pan (✉)
Brain Science Institute, Tamagawa University, Tokyo, Japan
e-mail: xcpan@lab.tamagawa.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_18,

parietal cortex [3, 4], and brain areas involved in motivation and emotion, such as the orbitofrontal cortex, the amygdala [5]. The LPFC sends outputs to motor-related areas [6]. The widely distributed connections allow the LPFC to integrate sensory information from various sources and emotional information to generate behavioral meaning of external stimuli. Second, electrophysiological studies in monkeys demonstrate that LPFC neurons encode category information for a group of stimuli with perceptual or functional similarity [7, 8]. Third, fMRI studies reported that the human prefrontal cortex responds to category learning and judgment [9, 10]. Taking together, we hypothesize that the LPFC may contribute to category inference.

To test this hypothesis, we recorded single-unit activity from the LPFC of two monkeys in a sequential paired-association task with an asymmetric reward schedule. We found that a group of LPFC neurons simultaneously encoded both the category information of a group of stimuli and reward information, and predicted reward for category members on the basis of category membership. These LPFC neurons might reflect a neural mechanism of category inference.

2 Results

We trained two monkeys to learn two association sequences (e.g. A1→B1→C1 and A2→B2→C2, Fig. 1a) in a sequential paired-association task [11]. After the

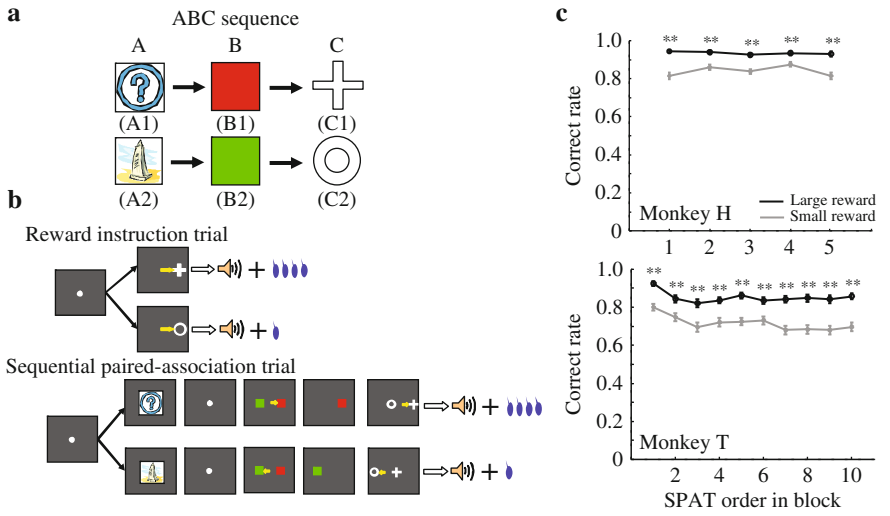


Fig. 1 The scheme of the task and the monkeys' behavioral performance. **a** Two associative sequences (ABC sequence) learned by monkeys. **b** An asymmetric stimulus-reward contingency was introduced in reward instruction trials, and used in the following SPATs in one block. **c** The correct rates of the first saccadic choice in SPATs for two monkeys (H and T). The horizontal axis indicates the SPAT order after reward instruction in one block. The *black curves* represent the correct rate in large reward trials, while the *gray curves* indicate the correct rate in small reward trials. $^{**}p < 0.01$ (two-tailed *t*-test). Error bars: s.e.m.

completion of learning, the monkeys were taught the asymmetric reward rule using reward instruction trials, in which one stimulus (C1 or C2) was paired with a large reward (0.4 ml) and the another stimulus (C2 or C1) with a small reward (0.1 ml). We arranged reward instruction trials and sequential paired-association trials (SPATs) in one block (Fig. 1b). At the beginning of each block, reward instruction trials were presented; then followed by SPATs with the same asymmetric reward rule. The stimulus-reward contingency was pseudo-randomized between blocks. Our question is whether the monkeys transfer the reward information associated with C1 and C2 to the first visual stimuli A1 and A2 in the SPATs.

The behavior of two monkeys was systematically influenced by the amount of reward. The correct rate of the first choice (selection of B on the basis of A) in the SPATs is significantly higher in large reward trials than in small ones from the first SPATs after reward instruction (Fig. 1c), indicating the monkeys transferred reward information associated with C1 and C2 to the stimuli A1 and A2.

We observed that a subpopulation of LPFC neurons predict reward information specific to a group of relevant visual stimuli (e.g. A1-group includes A1, B1, C1 and A2-group includes A2, B2 and C2) (Fig. 2). These neurons responded to each stimulus from the preferred group (A1-or A2-group) in the preferred reward condition (large or small), and showed no response to the stimuli from the non-preferred group irrespective of the reward condition. Thus, these neurons (category-reward neurons) likely encode both the category information of visual stimuli and reward information. The population activity of these neurons discriminated between preferred reward condition and non-preferred reward condition from the first SPATs in the condition of the preferred group of stimuli, but not of the non-preferred group of stimuli. The results suggest that LPFC neurons represent category-based reward information, and relay this information from one category member to others mediated by the common activity pattern of category-reward neurons to stimuli in the same category, which could be the neural basis of category inference.

One prediction from category inference is that LPFC neurons should transfer reward information from C1 or C2 not only to those well experienced stimuli (A1, A2, B1 and B2, Fig. 2c), but also to new category members that have never been used in the asymmetric reward task. To test this prediction, the monkeys were trained to learn associations between new visual stimuli and B1 or B2 (e.g. one pair of new stimuli (N1 and N2), N1→B1 and N2→B2). During the learning of new associations, the monkeys received the same amount of reward for all new stimuli. After completion of training, the monkeys acquired the new associations, but no asymmetric reward information paired with new stimuli. In the process of learning new associations, the monkeys categorized new stimuli into the current A1-group (or A2-group) through the association with common stimulus B1 (or B2). During the asymmetric reward task, these new stimuli were presented as the first visual cues in SPATs, instead of A1 and A2. We hypothesized that category-reward neurons similarly represent the category information for both the original (old) stimuli and the new stimuli in the same group, and transfer reward information to new stimuli even before these new stimuli have been associated directly with the asymmetric reward.

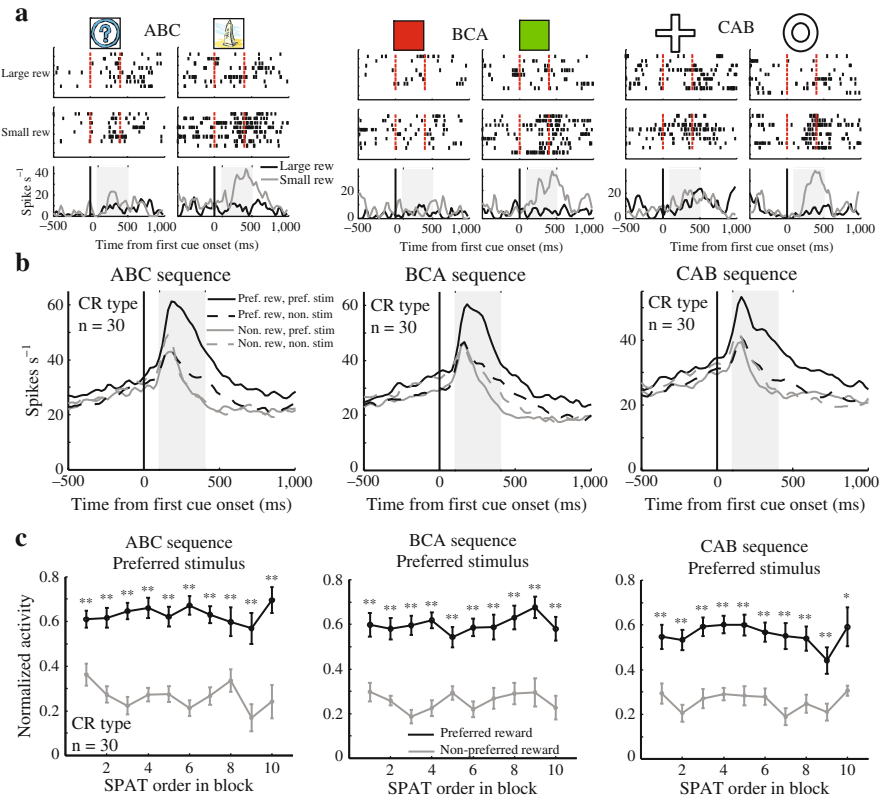


Fig. 2 A typical category-reward (CR) neuron and population activities in three sequences. **a** Single-unit activity of a CR neuron for three sequences: ABC (first cue: A1 or A2), BCA (first cue: B1 or B2) and CAB (first cue: C1 or C2) sequences. This neuron showed higher activity in the small reward trials than in the large reward trials when A2, B2 and C2 were presented as first cues, but showed no reward-modulated activity for A1, B1 and C1. Thus this neuron appears to predict reward in the condition of the preferred group of stimuli. The stimuli A2, B2 and C2 are referred to the preferred stimuli for this neuron. And the small reward condition in which the neuron produced higher activity is referred to the preferred reward condition. **b** The population histogram of CR neurons for three sequences. The activity for the preferred stimulus in the preferred reward condition was maintained across three sequences. **c** The normalized population activity of CR neurons for preferred stimuli in three sequences. Statistical significance was checked by two tailed *t*-test ($^{*}p < 0.05$, $^{**}p < 0.01$). Error bars: s.e.m.

Behaviorally, the correct rate of the first choice for new stimuli in the first presentation was significantly higher in large reward trials than in small ones. Category-reward neurons showed similar response patterns to both the old and the new stimuli, and predicted the reward information for the first presentation of new stimuli. These results suggest that LPCF neurons use category membership to transfer reward information to new members, consistent with the prediction from category inference.

3 Discussion

We observed category-reward neurons in the LPFC encoded the category information for the relevant stimuli, including old and new ones, in the preferred reward condition. We also found another group of neurons (called reward neurons) that showed reward modulated activity independent of visual stimuli, which could be a type of neurons observed in previous studies [12, 13]. Both types of neurons discriminated between the two reward conditions based on old or new stimuli from the first SPATs after reward instruction of C1 and C2, indicating these neurons can transfer reward information associated with C1 and C2 to old or new stimuli. The monkeys could not simply used mnemonic representations to predict reward for new stimuli. Effectively they had to integrate several independently acquired associations in order to infer reward value to new stimuli.

Category-reward neurons preferred one group of stimuli over the other. At the same time, they showed higher activity in the preferred reward condition than in the non-preferred one for the preferred stimuli, indicating that category-reward neurons encode both the category and reward information simultaneously. Category-reward neurons can predict reward information on the basis of the first presentation of new stimuli, once another stimulus from the same group was paired with large (or small) reward. Together, the current data suggest that category-reward neurons encode category-based reward information and propagate this information among stimuli in the same group, which would be the neural basis for category inference.

References

1. Jitsumori, M., Shimada, N., Inoue, S.: Family resemblances facilitate formation and expansion of functional equivalence classes in pigeons. Learn Behav. **34** (2006) 162–175.
2. Murphy, R.A., Mondragón, E., Murphy, V.A.: Rule learning by rats. Science. **319** (2008) 1849–1851.
3. Petrides, M., Pandya, D.N.: Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur. J. Neurosci. **11** (1999) 1011–1036.
4. Petrides, M., Pandya, D.N.: Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur. J. Neurosci. **16** (2002) 291–310.
5. Amaral, D.G., Price, J.L.: Amygdalo-cortical projections in the monkey (Macaca fascicularis). J. Comp. Neurol. **230** (1984) 465–496.
6. Miyachi, S., Lu, X., Inoue, S., Iwasaki, T., Koike, S., Nambu, A., Takada, M.: Organization of multisynaptic inputs from prefrontal cortex to primary motor cortex as revealed by retrograde transneuronal transport of rabies virus. J. Neurosci. **25** (2005) 2547–2556.
7. Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K.: Categorical representation of visual stimuli in the primate prefrontal cortex. Science. **291** (2001) 312–316.
8. Shima, K., Isoda, M., Mushiake, H., Tanji, J.: Categorization of behavioural sequences in the prefrontal cortex. Nature. **445** (2007) 315–318.
9. Vogels, R., Sary, G., Dupont, P., Orban, G.A.: Human brain regions involved in visual categorization. Neuroimage. **16** (2002) 401–414.
10. Jiang, X., et al.: Categorization training results in shape and category-selective human neural plasticity. Neuron. **53** (2007) 891–903.

11. Pan, X., Sawa, K., Tsuda, I., Tsukada, M., Sakagami, M.: Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. Nat. Neurosci. **11** (2008) 703–712.
12. Watanabe, M.: Reward expectancy in primate prefrontal neurons. Nature. **382** (1996) 629–632.
13. Kobayashi, S., Lauwereyns, J., Koizumi, M., Sakagami, M., Hikosaka, O.: Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. J. Neurophysiol. **87** (2002) 1488–1498.

Is Mu Rhythm an Index of the Human Mirror Neuron System? A Study of Simultaneous fMRI and EEG

Hiroaki Mizuhara and Toshio Inui

Abstract EEG mu rhythm is one of the prominent oscillatory patterns for investigating the dynamics of the human mirror system. However, because of the blurring effect caused by the volume conduction in EEG measures, the origin of this rhythm remains an open question. Here we propose a novel method to identify the EEG distribution on the scalp by using the simultaneous fMRI and EEG recording. The results indicate that the mu rhythm appeared on the lateral central sites indexes as activities in the sensorimotor cortex, while the contamination from activities in other cortices also appears as the mu rhythm.

Keywords fMRI · EEG · Mirror neuron system · Mu rhythm · Brain computer interface

1 Introduction

The mirror neuron system is one of the prominent neuronal correlates for emerging action understandings in human–human interaction. This system has often been investigated by using functional magnetic resonance imaging (fMRI), while the first report originally showed the existence of the mirror neuron to be in the motor cortex using a non-human primate study [1]. However, in order to investigate the dynamic property of the mirror neuron system during the human-human interaction, fMRI fundamentally lacks the ability to describe it because of the temporal resolution limits of the hemodynamic measures [2]. Therefore, the neuronal dynamics of the mirror neuron system in humans has sometimes been investigated by human electrophysiology (e.g., scalp electroencephalography: EEG). In the human scalp EEG, the mu rhythm (8–12 Hz, sometimes referred to as the rolandic alpha/beta rhythm) is used as an index of activation in the mirror neuron system [3, 4]. This is

H. Mizuhara (✉)
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
e-mail: hmizu@i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_19,

experimental evidence that the mu rhythm suppression appears at the lateral central sites on the scalp, which corresponds to the location just above the sensorimotor cortex, when humans are performing motor execution or motor imagery [5]. However, because of the blurring effects caused by the volume conduction when only by using the human scalp EEG, it is hard to identify the exact origin of this rhythm in the human brain. Therefore, it raises a question as to whether the scalp EEG mu rhythm truly originates from the sensorimotor cortex, or if it is the result of contamination from activities in other cortices.

In order to investigate if when the mu rhythm appeared on the scalp is an indication of the activity in the sensorimotor cortex, we measured the simultaneous fMRI and EEG during a motor execution task. The fMRI can provide detailed information of the location of activation. Therefore, the combination of the scalp EEG and fMRI could be a promising method to uncover the dynamics of the mirror neuron system. In the previous studies of the simultaneous fMRI and EEG recording, the cortical areas showing the significant correlation to the EEG time series have been identified by using the time series of the scalp EEG power or other indices as an event for the analyses [6, 7]. In the previous studies, since the blurring effect still remains in the EEG indices, it did not directly indicate if the EEG originated from a specific cortical area. To overcome this problem, we propose a new method for integrating the fMRI and EEG. The results directly show the cortical origin of the EEG activities observed on the scalp. Based on this new method, we further show that the scalp EEG mu rhythm reflects the activity in the sensorimotor cortex.

2 Materials and Methods

In the current study, 15 healthy males participated in the experiments with after giving written informed consent. The experimental procedure was approved by the ethics committee at Kyoto University. The task was the right/left hand grasping in accordance with a visually presented cue. Prior to the cues for hand grasping, a short pre-cue was presented for the purpose of shifting the participant's attention to the spatial location of the cue appearance. The cue for the right or left hand grasping appeared at almost the same frequency. During this hand grasping task, fMRI and EEG were simultaneously measured. For the fMRI data acquisition, a 1.5T MR scanner (Shimadzu-Marconi) was used with the following scan parameters: TR = 3 s, TE = 49 ms, FA = 90°, FOV = 192 mm, in-plane resolution = 64 × 64, and slice thickness = 5 mm without gap. The scalp EEG was acquired with 64ch EEG amplifier (Brainproducts) with the following parameters: sampling rate = 5 kHz, Highpass filter = 1 Hz, Lowpass filter = 250 Hz, Ref. Ch. = FPz, and Grd. Ch. = inion.

Figure 1 shows the flowchart of the proposed method to identify the scalp EEG topography regarding the cortical activity. The EEG data simultaneously measured with fMRI were contaminated by the scan and pulse artifacts caused by the high magnetic field. In the current study, these artifacts were removed by using the methods proposed by Allen et al. [8, 9]. The artifact removed EEG time series were then decomposed into time-frequency representation of EEG power by using the

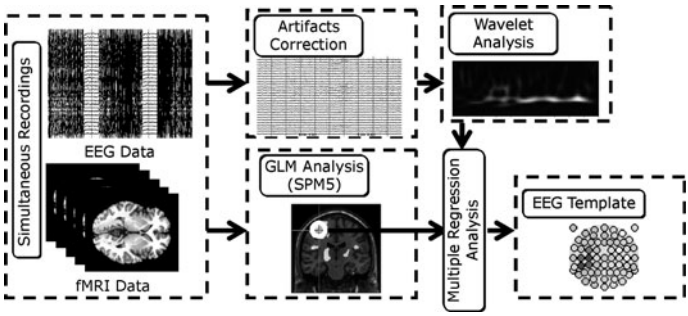


Fig. 1 Flowchart for identifying the scalp EEG topography related to the cortical activity based on the simultaneous fMRI and EEG

Morlet wavelet transformation, where the frequency ranged from 1 to 50 Hz with 30 steps in log-scale, and the sampling rate was down-sampled into 200 Hz prior to the wavelet transformation.

In the fMRI analysis, we used the SPM5 software (Wellcome Trust Centre for Neuroimaging). In this analysis, the fMRI data was preprocessed in the following order: motion correction to the first volume, slice timing correction to the middle slice of the entire scans, realignment to the 3D structure image, normalization to the MNI standard brain, and smoothing with the 10 mm FWHM Gaussian kernel. The event design was decided by the convolution of the hemodynamic response function (HRF) and the timing of the cue presentation. The voxel by voxel linear regression analysis among the preprocessed fMRI data and the event design was then performed to identify the cortical areas corresponding to the hand grasping. Based on the fMRI results, the regions of interest (ROIs) were then decided. It should be noted that the individual coordinate of the ROIs was decided by the individual fMRI results and structure MRI, and a sphere with a 5 mm radius surrounding the coordinate was decided for each individual participant.

For combining the fMRI and EEG data, we used the multiple regression analysis among the fMRI time series of ROIs and the temporal representation of EEG power. The fMRI time series were used for the multiple regressors, and the EEG power at each electrode and frequency was used as a regressand. Thus, the regression coefficient for the EEG power at each electrode represents the electric field from the neuronal activity in the ROI. Since the fMRI is based on the blood oxygenation level dependent signal, the fMRI signals respond with a time delay defined by the HRF [2]. Then, in the actual computation of multiple regression analysis, the regressand was transformed by the convolution of the EEG power time series and HRF.

3 Results and Discussion

Figure 2 shows the results of the EEG distributions corresponding to the activities in the left sensorimotor cortex, the left visual cortex and the left putamen during the hand grasping task. The EEG power (11.8 Hz) at the left lateral central electrodes

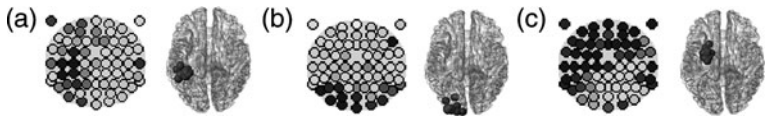


Fig. 2 Results of scalp EEG which correlated to the fMRI BOLD signal in the left sensorimotor cortex (**a**), the left visual cortex (**b**) and the left putamen (**c**). The figures appearing in the *left column* represent the EEG topography and the figures in the *right column* represent the corresponded cortices in each panel

showed the negative correlation to the BOLD signals in the left sensorimotor cortex. This indicates that the mu rhythm suppression occurs in association with the sensorimotor cortical activation [5]. In addition to this EEG suppression, EEG power at the left occipital electrode sites (9.5 Hz) and the left fronto-central sites (9.5 Hz) also showed negative correlation to the activation in the left visual cortex and in the left putamen, respectively. Previous studies reported that the hemodynamic response in the visual cortex negatively correlated to the EEG alpha power at the scalp occipital sites [10]. Furthermore, the mu rhythm was also found in the putamen during motor execution task [11]. These indicate that a part of mu rhythm can be interpreted as contamination from the activities in the visual cortex and putamen, while the mu rhythm also indexes the sensorimotor cortical activation in line with previous studies. Thus, our results strongly claim that one must carefully note the contamination from the activities in other cortices, when investigating the dynamics of mirror neuron system by using the scalp EEG.

Acknowledgments The authors express their appreciation to Dr S. Takano and Mr I. Fujimoto at ATR-Promotions, Japan for their technical support in the simultaneous fMRI and EEG. This study was partially supported by MIAC SCOPE-R (062108001), MEXT KAKENHI (20220003 and 21680024) and The Mitsubishi Foundation.

References

1. Rizzolatti, G., Fabbri-Destro, M.: The mirror system and its role in social cognition. Curr. Opin. Neurobiol. **18** (2008) 179–184.
2. Logothetis, N.K., et al.: Neurophysiological investigation of the basis of the fMRI signal. Nature. **412** (2001) 150–157.
3. Orgs, G., et al.: Expertise in dance modulates alpha/beta event-related desynchronization during action observation. Eur. J. Neurosci. **27** (2008) 3380–3384.
4. Ulloa, E.R., Pineda, J.A.: Recognition of point-light biological motion: mu rhythms and mirror neuron activity. Behav. Brain Res. **183** (2007) 188–194.
5. Wolpaw, J.R., et al.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113** (2002) 767–791.
6. Mizuhara, H., et al.: A long-range cortical network emerging with theta oscillation in a mental task. Neuroreport. **15** (2004) 1233–1238.
7. Mizuhara, H., Yamaguchi, Y.: Human cortical circuits for central executive function emerge by theta phase synchronization. Neuroimage. **36** (2007) 232–244.
8. Allen, P.J., et al.: Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. Neuroimage. **8** (1998) 229–239.

9. Allen, P.J., Josephs, O., Turner, R.: A method for removing imaging artifact from continuous EEG recorded during functional MRI. Neuroimage. **12** (2000) 230–239.
10. Goldman, R.I., et al.: Simultaneous EEG and fMRI of the alpha rhythm. Neuroreport. **13** (2002) 2487–2492.
11. Sochurkova, D., Rektor, I.: Event-related desynchronization/synchronization in the putamen: an SEEG case study. Exp. Brain Res. **149** (2003) 401–404.

Decoding Action Selectivity of Observed Images Using fMRI Pattern Analysis

Kenji Ogawa and Toshio Inui

Abstract Previous research indicates that mirror neuron system (MNS), comprising posterior parietal (PPC) and ventral premotor (PMv) cortices, has a role in action understanding. However, action selectivity within these regions has not been elucidated. We investigated whether the MNS contains information about observed actions. Participants observed object-directed hand actions. We independently manipulated five dimensions of observed images: action (grasp/touch), object (cup/bottle), perspective (egocentric/allocentric), hand (right/left), and image size (large/small), and investigated whether this information could be decoded from multi-voxel pattern activity. The early visual area showed significant above-chance classification accuracy consistent with dissimilarity of input images. In contrast, significant decoding accuracy was observed for action, object, and mirror-image of hand in the PPC and for action and object in the PMv. Our study indicates that the MNS processes observed actions in a hierarchical manner, where the PPC represents action together with hand, and the PMv encodes more abstracted representations of transitive action.

Keywords Action observation · Posterior parietal cortex · Premotor cortex · fMRI

1 Introduction

Humans can recognize observed actions despite variations in retinal inputs such as viewpoint and distance from others as well as the hand used, which requires invariant representations of observed actions. Recent research indicates that the human mirror neuron system (MNS), or action observation network, mainly comprising the posterior parietal cortex (PPC) and the ventral premotor region (PMv), has a role in the understanding of action [1]. However, most previous fMRI studies have failed to reveal action selectivity within these regions due to its limited spatial

K. Ogawa (✉)
ERATO Asada project, Japan Science and Technology Agency, Kyoto 606-8501, Japan
e-mail: ogawa@cog.ist.i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_20,

resolution, which is most critical for action understanding. In this study, we used recently-developed multi-voxel pattern analysis (MVPA) over fMRI data [2, 3] to investigate neural representations of observed actions in the parieto-premotor area, regarding whether these areas contain specific information about observed actions.

2 Materials and Methods

Nine participants (5 males and 4 females; mean 29.0-year old; one left-handed) observed grayscale photos of object-directed hand actions. We independently manipulated the following five properties of the observed images: action (grasp or touch), manipulated object (cup or bottle), perspective (egocentric or allocentric), hand (right or left), and image size (large or small), resulting in a total of 32 types of stimuli (Fig. 1). These were presented in separate blocks per run. This manipulation enabled us to investigate distinct neural representations encoded in different brain regions, as well as to dissociate low-level visual dissimilarity with action selectivity. Ten images of the same properties were presented every 1,200 ms, with 800 ms for image presentation and 400 ms for a fixation display during each 12-s block. Every block was interleaved with a 12-s rest period. To ensure the subject's attention during scanning, we introduced subtle variations into the images, in which the hand's shape or orientation was slightly changed across images. Participants performed a one-back task, simultaneously pressing buttons with index fingers of both hands whenever the same stimulus appeared twice in a row, which occurred once or twice per block.

A 3-T scanner (Simens, Trio) was used to acquire T2*-weighted echo planar imaging with resolution of a 3 mm-cube. Image preprocessing was performed with SPM5, and classification analysis was performed with a linear support vector machine (SVM). The 32 blocks (trials) were modeled as separate box-car regressors that were convolved with a canonical hemodynamic response function. Parameter estimates of each trial of voxels within regions of interest (ROIs) were used as inputs to the classifier. Voxels in each ROI were selected in the order of highest *t*-value of training runs until the number of voxels reached 300 for each subject. The ROIs included the anterior part of intraparietal sulcus (aIPS) and the ventral premotor

Fig. 1 The sample stimuli used in the experiment. (*Left*) *Grasp* the *cup* with *right hand* in *egocentric perspective* displayed in *small size*. (*Right*) *Touch* the *bottle* with *left hand* in *allocentric perspective* displayed in *large size*

cortex (PMv) bilaterally. These were defined as 15 mm-radius spheres centered at the mean coordinates from previous literature of hand-action execution and observation studies listed in [4, 5]. We also chose the right early visual area (BA17/18, Vis) defined with the SPM Anatomy toolbox, as well as the white matter (WM) as a control ROI. We then evaluated selectivity as well as invariance to different properties (dimensions) of the observed images, focusing on whether this information could be correctly decoded from multi-voxel pattern activity of the region (information-based analysis proposed in [6]). Hand was labeled in two ways: either as anatomical- (right hand in allocentric corresponds to right hand in egocentric perspective) or mirror- (right-hand in allocentric corresponds to left-hand in egocentric perspective) correspondence. Decoding accuracy was measured with 4fourfold "leave-one-out" cross-validation based on independent runs, and statistical significance was assessed with a t test (df = 8).

3 Results

The conventional univariate analysis of fMRI data showed bilateral large activations during task blocks, compared with rest periods, collapsed over all conditions, which encompassed parieto-premotor region including the aIPS and PMv, as well as the early visual areas and medial motor areas including supplementary and pre-supplementary motor areas (Fig. 2). This result is generally congruent with previous studies of action observation. MVPA revealed that the early visual area showed significant above-chance classification accuracy in all dimensions, with particularly high accuracy in image size, hand, and perspective. These patterns of decoding accuracy were mostly consistent with the dissimilarity of the low-level stimulus images, which was measured with pixel-wise correlations as well as decoding accuracies using a stimulus image itself as input. In contrast, significant above-chance decoding accuracy was observed in the bilateral aIPS, showing highest accuracy in action as well as object and mirror-image hand, with no significant accuracy for size and anatomical-hand. The bilateral PMv showed significant accuracy for action, together with object selectivity in the right PMv, without significant accuracy for size, hand, and perspective (Fig. 3).

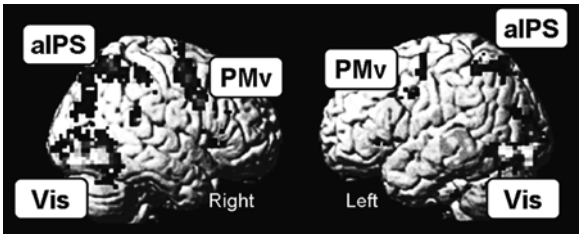


Fig. 2 Overall activated regions in an action observation period compared with a rest period

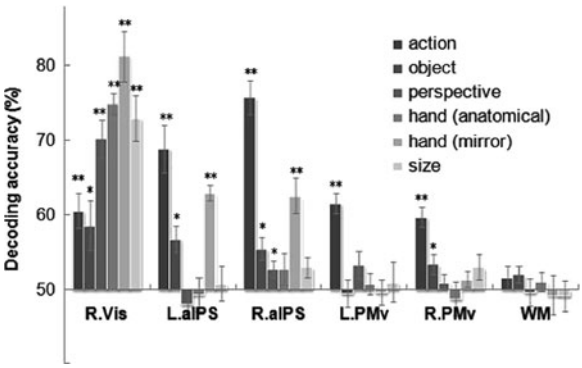


Fig. 3 Decoding accuracies of regions of interests (ROIs) for six different properties of observed images. Error bars denote SEMs across subjects. *, $p < 0.05$; **, $p < 0.005$

4 Discussion

Our results revealed that while the early visual areas process low-level visual information of input stimulus, the parieto-premotor areas that comprise the MNS encode action-dominant information regarding observed images. The MNS appears to processes observed actions in a hierarchical manner, where aIPS represents object-directed action together with hand used, and PMv encodes a more abstracted representation of transitive actions invariant to the effector. The current results are consistent with those of previous studies that indicated action-related activity of the PMv that was invariant to effectors and responded to actions made with tools as well as hand and mouth [7], or to action-related sound [8]. In contrast, the PPC encodes effector selectivity of observed action [9]. These different neural representations of observed actions within the parieto-premotor cortex, as well as early visual areas, may be responsible for different levels of action understanding as well as different capacities for imitation.

Finally, higher decoding accuracy for action was shown in the parietal cortex (aIPS) than in the early visual areas (Fig. 3), despite the fact that we used only visual stimuli. Previous decoding studies mostly targeted the primary visual areas to decode observed images (e.g. [3]). In contrast, our study suggests the merit of using activation patterns of the higher brain regions, rather than the primary visual areas, to decode invariant information (e.g. action selectivity in the current study) that is distinct from low-level visual similarity of input images.

References

1. Rizzolatti, G., Fogassi, L., Gallese, V.: Neurophysiological mechanisms underlying the understanding and imitation of action. Nat. Rev. Neurosci. **2** (2001) 661–670.
2. Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V.: Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. **10** (2006) 424–430.
3. Kamitani, Y., Tong, F.: Decoding the visual and subjective contents of the human brain. Nat. Neurosci. **8** (2005) 679–685.

4. Chong, T.T., Cunnington, R., Williams, M.A., Kanwisher, N., Mattingley, J.B.: fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Curr. Biol. **18** (2008) 1576–1580.
5. Dinstein, I., Gardner, J.L., Jazayeri, M., Heeger, D.J.: Executed and observed movements have different distributed representations in human aIPS. J. Neurosci. **28** (2008) 11231–11239.
6. Kriegeskorte, N., Goebel, R., Bandettini, P.: Information-based functional brain mapping. Proc. Natl. Acad. Sci. U. S. A. **103** (2006) 3863–3868.
7. Ferrari, P.F., Rozzi, S., Fogassi, L.: Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. J. Cogn. Neurosci. **17** (2005) 212–226.
8. Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., Rizzolatti, G.: Hearing sounds, understanding actions: action representation in mirror neurons. Science. **297** (2002) 846–848.
9. Shmuelof, L., Zohary, E.: Mirror-image representation of action in the anterior parietal cortex. Nat. Neurosci. **11** (2008) 1267–1269.

A Developmental Model of Infant Reaching Movement: Acquisition of Internal Visuomotor Transformations

Naohiro Takemura and Toshio Inui

Abstract The online control mechanisms of human goal-directed movements include two functions: predicting sensory feedback resulting from motor commands (forward transformation) and generating motor commands based on sensory inputs (inverse transformation). Infants acquire these transformations without implicit instructions from caregivers. In this paper, a neural network model is proposed that learns the forward and inverse transformations in reaching movement by observing the randomly moving hand. The forward pathway of the model is a Jordan network that is input with motor commands, and that is trained to output a visual hand position. The inverse pathway has the input of the visual hand position and a connection from the hidden layer of the forward pathway. It is trained to output the motor command, which makes the hand move to the input hand position. The network learned correct transformations, which suggests that continuous observation of the hand is the basis for motor development.

Keywords Infant development · Reaching · Neural network model

1 Introduction

The internal model has been suggested to have an important role in the control of goal-directed human movements. In order to achieve the online control of reaching towards a visually presented target, the central nervous system has to compensate for the intervention of biological delay and noise in motor and sensory pathways. To this end, the internal forward model of the arm predicts the body state, as well as the sensory feedback that is the consequence of the motor command. On the other hand, the inverse model computes the required motor command from the predicted

N. Takemura (✉)
JST ERATO Asada Synergistic Intelligence Project, Kyoto 606-8224, Japan
e-mail: takemura@jeap.org

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_21,

body state. An optimal feedback control model based on these forward and inverse transformations has been suggested as a framework for online motor control [1].

These motor control mechanisms are typically acquired during infancy. Infants acquire the capacity of motor prediction using the forward model by 7 months after birth [2]. However, it is still not clear how infants learn the forward and inverse transformations without implicit instructions from their caregivers. On the other hand, infants have tendency to keep their hands within their field of view [3]. In this paper, a neural network model is presented that learns the forward and inverse transformations by observing a hand that moves randomly within the visual range.

2 The Model

The model learns the forward and inverse transformations of a hand that moves in 2 dimensional space, through the visual feedback of the hand position. In the current study, the motor command is the hand displacement from current position, which is, during learning, issued by random sampling from an uniform distribution ranging from -2 to 2 for each dimension. The visual feedback is the hand position in that space, ranging from -5 to 5 for each dimension. Following an infant's tendency to keep the hand within view, the hand position is assumed to be always within the visual range. When a motor command that moves the hand out of the visual range is issued, the hand stops at the end of the visual range.

The model consists of two pathways, the forward transformation pathway and the inverse transformation pathway (Fig. 1). The forward transformation pathway is a Jordan network, where motor commands are input, and is trained with a backpropagation algorithm. The teacher signal is the visual feedback of the hand position as a consequence of the issued motor command. On the other hand, the inverse transformation pathway is a network where the input is the visual feedback of the hand position. It also has connection from the output of the hidden layer of the forward pathway from 1 time step previous. The network is trained with the backpropagation. The motor command that causes the hand position input through the visual feedback is the teacher for training.

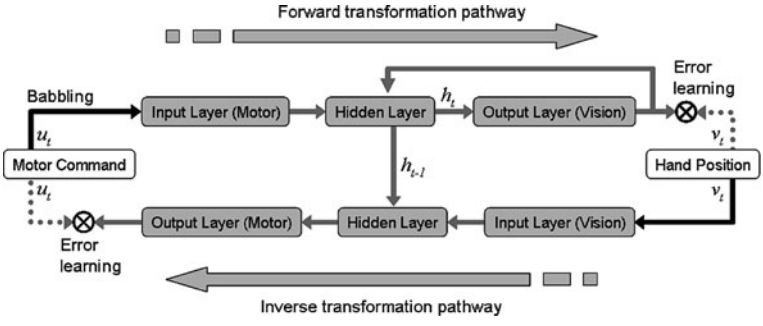


Fig. 1 Schematic diagram of the model

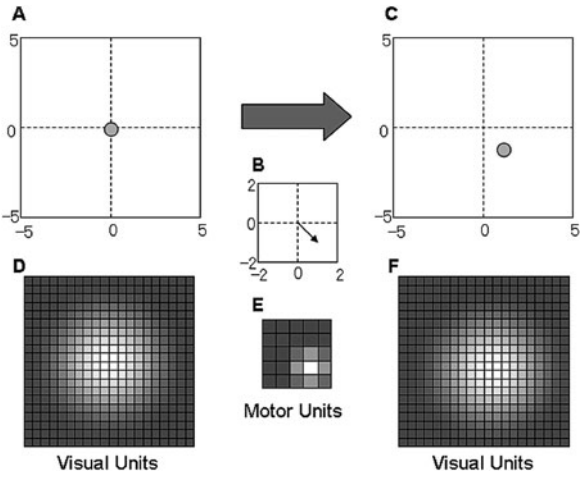


Fig. 2 Examples of the population activity of the motor and visual units. Given that the current hand position is at the center of the visual range (**a**: *circle*), and that the motor command (**b**: *arrow*) is issued to make the hand displacement of (1,−1), the resultant hand position is (1,−1) (**c**: *circle*). The hand positions and the hand displacement are encoded in the population activity (**d, f**, and **e**), where the *gray scale color* of each cell denotes the activity level (the more *white*, the higher the activity), and the position of each cell in the population denotes the unit's preference for the corresponding hand positions and hand displacement

The motor command and the visual feedback are encoded by a population of units that have broad selectivity for hand displacement and hand position, respectively. The selectivity of each unit is cosine tuned; that is, given that the hand displacement or hand position to encode is x, the activation a_i of a unit i, which has preferential hand displacement or hand position p_i, is defined as follows:

$$a_i = \begin{cases} 0.5\cos(|x - p_i|/2W) + 0.5 & \text{if } |x - p_i| < W \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Note that $|\cdot|$ denotes norm, and W is half the size of the motor and visual range (i.e. 2 and 5 for the hand displacement and the hand position, respectively). The number of the motor units (i.e. the input to the forward pathway and the output of the inverse pathway), the vision units (i.e. the output of the forward pathway and the input to the inverse pathway), and the hidden units are 5×5, 20×20, and 500, respectively. Examples of representations of the visual feedback and hand position are shown in Fig. 2.

3 Simulation and Discussion

The network is trained in 300,000 steps with the pseudo-random motor command. After learning, 25 motor commands are input to the forward transformation

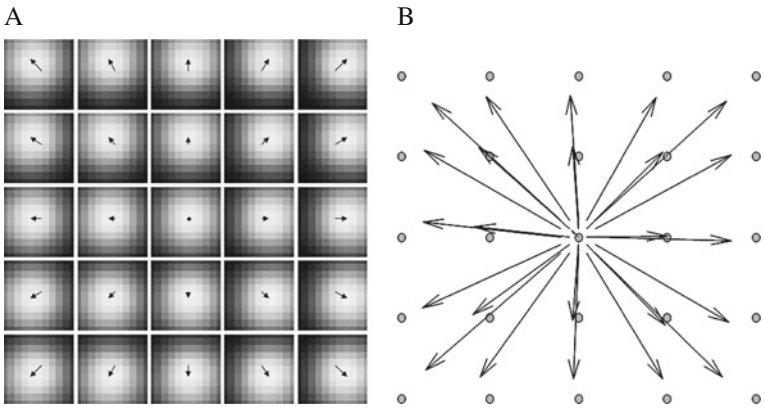


Fig. 3 Results of a network test. **a** Population Activities (*gray scale maps*: the more white, the higher the activity) of output units in the forward pathway when 25 motor commands (*arrows*) are input to the network, with the initial hand position being the center of the visual range. **b** Decoded hand displacements (*arrows*) from the population activities of output units in the inverse pathway when 25 new virtual hand positions (*circle*) are input to the network, with the initial hand position being the center of the visual range

pathway, with the initial hand position being the center of the visual range. The activities of output visual units are shown in Fig. 3a. The center of activation is almost the hand position after movement. On the other hand, 25 new virtual hand positions (ranging from (−2,−2) to (2,2)) are input to the inverse transformation pathway, with the initial hand position being the center of the visual range. The population activities of motor output units represent the motor commands that make the hand move to the input virtual hand position (Fig. 3b).

These results suggest that the forward and inverse transformations can be learned only by making random movement and by continuously looking at the moving hand. The correct inverse transformation indicates that the hidden layer units in the forward transformation contain the internal representation of the hand position. With these learned representations and transformations, infants may have the ability to acquire the online control mechanisms. In other words, an infant's tendency to keep the hand in the field of view is a behavior that makes motor development easier. One of the remaining problems is whether this tendency is an innate property or an acquired performance.

References

1. Todorov, E., Jordan, M.I.: Optimal feedback control as a theory of motor coordination. Nat. Neurosci. **5** (2002) 1225–1235.
2. Witherington, D.C.: The development of prospective grasping control between 5 and 7 months: A longitudinal study. Infancy **7** (2005) 143–161.
3. van der Meer, A.L.H., van der Weel, F.R., Lee, D.N.: The functional significance of arm movements in neonates. Science **267** (1995) 693–695.

Interacting Humans and the Dynamics of Their Social Brains

Emmanuelle Tognoli, Gonzalo C. de Guzman, and J.A. Scott Kelso

Abstract Social behavior rests on a complex spatio-temporal organization of brain processes in the interacting humans. We present a theoretical and methodological framework for the systematic study of brains and behaviors during social interactions. We draw an overview of results from our laboratory that describe human tendencies for coupled behaviors, the dynamics of their characteristic neuromarkers, and neuro-social phenomena in which one human can drive the brain dynamics of others or several humans can see their brain dynamics coupled, presumably by way of shared-attention to salient features of their joint-behavior.

Keywords Social behavior · EEG · Brain Coordination dynamics · 4D spatio-temporal analysis

1 Introduction

Two complementary forces of integration and segregation are at play in social behavior. Too much segregation and meaningful social interactions cannot emerge. On the other hand, too much integration prevents complex interactions – distinct yet coordinated behavior between partners –from arising. In between the two poles of integration and segregation lies meaningful social behaviors, the temporal unfolding of which requires a dynamical framework [1]. The overarching goal of Social Coordination Dynamics is to understand the mechanisms through which the behaviors and brains of people couple and decouple dynamically during the course of social interaction.

The basic experimental paradigm of Social Coordination Dynamics sets up subjects in reciprocal behavioral interactions [2] and continuously records state variables of their behavior along with simultaneous brain measures such as EEG [3].

E. Tognoli (✉)
The Human Brain and Behavior Laboratory, Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL 33431, USA
e-mail: tognoli@ccs.fau.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_22,

Brain processes associated with social interactions change at a fast sub-second time scale. Therefore, experimental paradigms are designed with the goal of knowing at every moment in time if the partners' behavior is coordinated or not. In coordination dynamics, the coordination states of a system evolve over time, as observed through special quantities called collective variables. In order to answer questions regarding the neural mechanisms responsible for establishment or loss of social coordination, it is crucial to understand how the continuous collective variables of behavior are related to the continuous dynamics of brains, when subjects are acting singly and together. To investigate the dynamics of coordination at behavioral and brain levels, a novel EEG approach was developed [4, 5]. This method reveals how patterns of brain activity within and between brains are associated with salient features of social behavior.

1.1 The Theoretical Framework of Coordination Dynamics

Human social systems are composed of people, their behaviors and their brains, a shared environment and information flowing between all these parts. Coordination Dynamics [6] aims to describe how such a complex system functions. It does so by studying how the parts' coordination evolves over time. An example of coordination dynamics is the dynamical coupling and uncoupling observed between the behavior of one person and brain activity in another person. Distinctive from other frameworks, Coordination Dynamics does not require that the parts that it studies be of the same nature (e.g. only brain regions or only people). Sufficient conditions are that the parts exchange information and can be described quantitatively over time. Relevant heterogeneous parts of systems are not isolated in separate scientific approaches or disciplines, thereby insight is enhanced.

1.2 Dynamics: Challenge for Social Neuroscience

The neural bases of social behavior are increasingly well-known [7, 8] especially in the wake of Rizzolatti and colleagues' discovery of specialized neurons of the motor cortex that share the dual property of acting for the self and perceiving social others. However, less is known about the dynamics of the brain during social interactions: how those networks of brain regions couple, decouple and re-arrange in new self-organized patterns of social interaction.

EEG and MEG are excellent techniques to study the large-scale coordination dynamics of the brain because of their wide coverage of the cortical areas and their fast temporal resolution. We recently developed specific methods to deal with continuous EEG activity that comes along spontaneously occurring events in the social coordination behavior [3–5]. One milestone in the development of this methodological framework was to get away from averaging techniques and address continuous brain activity [4]. Another was to design a visualization technique that simultaneously displays 4 dimensions of EEG data (2D of space, magnitude of the signal and time) [3, 5], thereby exposing the dynamic patterning of the brain [5]. Using

this method, it was possible to integrate the dynamics of behavioral and brain variables in a single analytical framework. In the following, we will first describe social behavior and then examine its associated brain coordination dynamics.

2 Self-organizing Coordination Tendencies in Human Behavior

2.1 Spontaneous Coordinative Tendencies

A basic model for self-organized behavior in a single person is the coordination between body parts: an essential property of human locomotion and action. Stable patterns of coordination have been comprehensively explored and modeled with the basic HKB equations of coordination dynamics and its many extensions [9]. As Rizzolatti and colleague showed [7], basic building blocks serving individual behaviors are also employed to acquire and use information about others' behavior. We studied the coordination of rhythmic movements between pairs of people who exchanged information about each others' behavior through visual perception. Coordination spontaneously occurred [2, 3]: the same attractors in the system dynamics that exist within subjects were observed between subjects (see also [10]).

2.2 Symmetry Breaking and the Notion of Roles

In spontaneous [3] and intentional social coordination tasks [11], close examination of the transition to synchronized behavior revealed that subjects are not identically involved in setting the collective behavior. Asymmetries emerge in the mechanisms and extent to which individual behaviors are modified. For instance, a subject may show little or no modification of his ongoing behavior and the partner may carry all the burden of establishing an instructed coordination pattern. Both subjects may adopt different roles, altering distinct parameters of their movements. Qualitatively, asymmetry is revealed by a detailed behavioral analysis of the dynamics of frequency, amplitude, and velocity of the interacting subjects' behavioral trajectories. Quantitatively, asymmetry may be analyzed through the directionality in the informational flows between the behaviors [12].

3 Brain Dynamics of Interacting Human

3.1 Neuromarkers of Social Coordination

During spontaneous social behaviors, EEG spectra revealed a neuromarker of social coordination, maximal above right centro-parietal electrodes, which we called phi [3]. Phi belongs to the 10 Hz frequency range and is composed of two components: a lower frequency component that increases during uncoordinated behavior and a higher frequency one that increases during sustained coordination.

3.2 Dynamics of Social Coordination

A time-frequency analysis showed that phi was not sustained or amplitude modulated. Rather, phi appeared intermittently, suggesting that much information about its specific function was hidden in its temporal distribution in relation with behavioral descriptors of the subjects' roles and circumstances. To study the temporal distribution of phi and other neuromarkers, EEG recorded during social interaction was filtered in the band of interest and spatio-temporal patterns were analyzed using a 4D method [5].

3.3 Brain Patterns of Coordination, Patterns of Intrinsic Behavior

Informationally, transitions are key episodes in the behavior of a system. With the aim of identifying dynamics of phi and other neuromarkers during social behavior, we focused on the transition to coordinated or uncoordinated behavior in a task of intentional social coordination. That is, on seeing the other's hand movements, subjects were instructed to intentionally coordinate in-phase or anti-phase with one another or to try to move independently by retaining their own intrinsically preferred behavior. In addition to the coordination or lack thereof, each transition was described with the roles taken by the subject in the outcome behaviors: intention to coordinate, mechanisms of change, congruency with the instructions and kinematic descriptors. The analysis identified the temporal distribution of phi and of other neuromarkers involved in social or intrinsic behavior. Phase of movement (flexion, extension) was an important determinant in the brain dynamics. Independently from the particular brain patterns occurring, there was a tendency for these brain patterns to change coincidently with switches in the behavioral phase of one of the partners (sometimes the self, sometimes the other). Perhaps as a result of saliency of one of the two behaviors, there was also a tendency for the transitions in brain patterns to be synchronized between the two partners that simultaneously engaged in the production of a collective behavior.

4 Discussion and Outlook

Study of continuous brain dynamics reveals a subtle interplay of integration and segregation during social behavior. Social coordination and other brain mechanisms such as social attention that may not have revealed their task-dependency in temporally coarse studies of average spectra have been clarified. Examination of the interaction between multiple neuromarkers has begun such as the coupling between phi and mu neuromarkers during behavioral changes aimed at coordination or apartness. Mechanisms of coupling between the brains of interacting partners have also been sketched.

Further insight will be gained in full parametric studies of relevant system variables. This aim will be assisted by a novel framework called Virtual Partner Interaction [13]. In this hybrid framework at the edge between modeling and

experiment, one of the human partners is replaced by a machine modeled after him/her, and the interaction can be set up to explore extended regions of the parameter space, with the intention to expose more fully the brain dynamics of the human partner.

Acknowledgments This work was supported by NIMH Innovations Award MH 42900, NIMH Grant MH80038, NSF Grant BCS0826897, the US Office of Naval Research Contract N000140510117. ET is supported by the Davimos Family Endowment for Excellence in Science. The contributions of Julien Lagarde, William McLean and Daniela Benites in various aspects of this research framework are acknowledged.

References

1. Kelso, J.A.S., Engstrøm, D.: The Complementary Nature. Cambridge: MIT Press (2006).
2. Oullier, O., de Guzman, G.C., Jantzen, K.J., Lagarde, J., Kelso, J.A.S.: Social coordination dynamics: measuring human bonding. Soc. Neurosci. **3** (2008) 178–192.
3. Tognoli, E., Lagarde, J., de Guzman, G.C., Kelso, J.A.S.: The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. U. S. A. **104** (2007) 8190–8195.
4. Tognoli, E.: EEG coordination dynamics: neuromarkers of social coordination. In Fuchs, A., Jirsa, V.K., eds.: Coordination: Neural, Behavioral and Social Dynamics. Berlin, Heidelberg, New York, NY: Springer-Verlag (2008) 309–323.
5. Tognoli, E., Kelso, J.A.S.: Brain coordination dynamics: true and false faces of phase synchrony and metastability. Prog. Neurobiol. **87** (2009) 31–40.
6. Kelso, J.A.S.: Dynamic patterns: the self-organization of brain and behavior. Cambridge: MIT Press (1995).
7. Rizzolatti, G., Craighero, L.: The mirror neuron system. Ann. Rev. Neurosci. **27** (2004) 169–192.
8. Ochsner, K.N., Lieberman, M.D.: The emergence of social cognitive neuroscience. Am. Psychol. **56**(9) (2001) 717–734.
9. Kelso, J.A.S.: Coordination dynamics. In Meyers, R.A., ed.: Encyclopedia of complexity and system science. Berlin, Heidelberg, New York, NY: Springer-Verlag (2009).
10. Schmidt, R.A., Richardson, M.J.: Dynamics of interpersonal coordination. In Fuchs, A., Jirsa, V.K., eds.: Coordination: Neural, Behavioral and Social Dynamics. Berlin, Heidelberg, New York, NY: Springer-Verlag (2008) 282–308.
11. Tognoli, E., Magne, C., de Guzman, G.C., Tuller, B., Kelso, J.A.S.: Brain Rhythms Underlying Intentional Social Coordination. [CDROM] Program No. 304.24, Abstract Viewer/Itinerary Planner. San Diego: Society for Neuroscience (2007).
12. De Guzman, G.C.: Fine structure of spontaneous visual social coordination. Proceedings of the GSA Symposium – Moving to Dynamical Framework in Aging Research – Visions and Tools for the New Frontier, Orlando, FL, 2005.
13. Kelso, J.A.S., de Guzman, G.C., Reveley, C., Tognoli, E.: Virtual partner interaction (VPI): exploring novel behaviors via coordination dynamics. PLoS One. **4**(6) (2009) e5749.

State-Dependent Cortical Synchronization Networks Revealed by TMS-EEG Recordings

Keiichi Kitajo, Ryohei Miyota, Masanori Shimono, Kentaro Yamanaka, and Yoko Yamaguchi

Abstract Transcranial magnetic stimulation (TMS) can noninvasively modulate cortical ongoing activity in the human brain. We investigated frequency-specific and state-dependent cortical network by analyzing how modulation of cortical ongoing activity at one cortical area is propagated to the rest of the brain by TMS-EEG recordings. We found frequency-specific and state-dependent changes in propagation of TMS-evoked phase resetting of cortical ongoing activity in the open eye condition and closed eye condition. We discussed the functional significance of state-dependent synchronization networks observed.

1 Introduction

Synchronization of oscillatory neural activity has been proposed as a fundamental mechanism underlying transient functional coupling of neural assemblies [1, 2]. Growing body of experimental evidence indicates that synchronization of neuronal oscillations plays an important role in dynamically linking multiple brain regions establishing information transfer and integrating distributed neural activity in the brain during cognitive and perceptual processing [3, 4]. In humans, brain-wide frequency-specific neural synchrony has been associated with perception of complex visual stimuli such as faces [5], perceptual awareness [6, 7], attention [8], and central executive function [9].

Transcranial magnetic stimulation (TMS) can transiently perturb and modulate cortical ongoing activity in the human brain. A previous study demonstrated evidence for a breakdown of long-range effective connectivity during NREM sleep by TMS–EEG (electroencephalography) recordings [10]. No study, however, has shown frequency-specific, state-dependent changes in large-scale cortical synchronous network connectivity by TMS–EEG. We therefore demonstrate

K. Kitajo (✉)
Laboratory for Dynamics of Emergent Intelligence, RIKEN Brain Science Institute, Saitama 351–0198, Japan
e-mail: kitajyo@p.u-tokyo.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_23,

frequency-specific and state-dependent cortical network connectivity by analyzing how TMS-evoked modulation of ongoing activity at one cortical area measured by EEG is propagated to the rest of the brain at different frequencies. Specifically, we investigated frequency-specific changes in propagation of TMS modulated ongoing activity in the open eye condition and closed eye condition.

2 Methods

Using a TMS-compatible 19 ch EEG amplifier (Neuroprax, neuroConn GmbH, Germany), we recorded EEG of ten participants (20–44 year, 7 males and 3 females) at a sampling rate of 4,096 Hz off line resampled at 1,024 Hz. Participants gave informed consent and sit on a chair with their eyes closed or eyes open fixating a cross on a display. Using a figure-of-eight coil connected to a biphasic stimulator (Magstim Rapid, Magstim Company Ltd, UK) single-pulse TMS were delivered to the left primary motor cortex at random intervals with intensity at the 95% motor threshold. The amount of auditory noise associated with each TMS pulse was minimized by ear plugs. Participants were given 50 TMS pulses for open eye and closed eye conditions.

We computed instantaneous phase and amplitude of the filtered EEG signals and calculated the phase locking values *PLV* as a measure for TMS-evoked phase resetting at each electrode at different frequencies. We first band-pass filtered the EEG recordings using the Morlet wavelet around central frequencies (cf) from 2 to 35 Hz in 0.2-Hz steps. Next, we defined the instantaneous phase of the filtered signal by constructing an analytic signal [11]. The analytic signal $\zeta(t)$ of an arbitrary signal $f(t)$ can be calculated as

$$\zeta(t) = f(t) + i\tilde{f}(t) = A(t)e^{i\phi(t)}, \tag{1}$$

where the function $\tilde{f}(t)$ is the Hilbert transform of $f(t)$,

$$\tilde{f}(t) = \pi^{-1}\text{P.V.}\int_{-\infty}^{\infty}\frac{f(\tau)}{t-\tau}d\tau, \tag{2}$$

and P.V. means that the integral is taken in the sense of Cauchy principal value. The instantaneous amplitude $A(t)$ and the instantaneous phase $\phi(t)$ of $f(t)$ are thus uniquely obtained from Eq. (1). We calculated the *PLV* between the instantaneous phases of filtered EEG records from different epochs as

$$PLV(t) = \frac{1}{N}\left|\sum_{n=1}^{N} e^{i\phi(t,n)}\right|, \tag{3}$$

where ϕ is the instantaneous phases of EEG records from an electrode, and N is the number of EEG epochs included in the calculation. *PLV* is a real value between 1 (maximum phase locking) and 0 (no phase locking). To detect TMS-related phase

resetting, standardized *PLV* values, $PLV_z(t)$, were computed from

$$PLV_z(t) = (PLV(t) - PLV_{\mathrm{Bmean}})/PLV_{\mathrm{Bsd}}, \quad (4)$$

where $\mathrm{PLV}_{\mathrm{Bmean}}$ and $\mathrm{PLV}_{\mathrm{Bsd}}$ are the mean and standard deviation of PLVs computed from the 400-ms pre-TMS baseline at each frequency. The resulting index, PLV_z, indicates standardized changes in phase locking across epochs. To assess statistical reliability of the PLV_z, the threshold was set as a function of the distribution of PLV_z during the resting baseline period.

3 Results and Discussion

Figure 1 shows averaged time-frequency diagrams of PLV_z for open eye, closed eye conditions and difference between them. TMS-evoked strong phase resetting of ongoing activity was observed at around 2–5 Hz in the open eye and closed eye conditions. We found differences in spatio-temporal patterns of propagation of phase resetting in open eye and closed eye conditions most prominently at 2–5 Hz. The phase resetting propagated from the left primary motor area to posterior and contralateral areas more prominently in the open eye condition than in the closed eye condition (Fig. 2).

The results indicate that state-dependent TMS-evoked neural transmission occurs across the distant brain regions in a frequency-specific way across coupled neural oscillators. The state-dependent changes in large-scale synchronization network connectivity in open eye and closed eye conditions might be related to changes in baseline network activity such as visual network activity in the posterior areas between the open eye and closed eye conditions. The state-dependent changes might be functionally associated with differences in perceptual set and/or attentional state between two conditions. Our study suggests we can investigate the dynamics of frequency-specific cortical and/or thalamocortical synchronization networks in the human brain by TMS-EEG recordings.

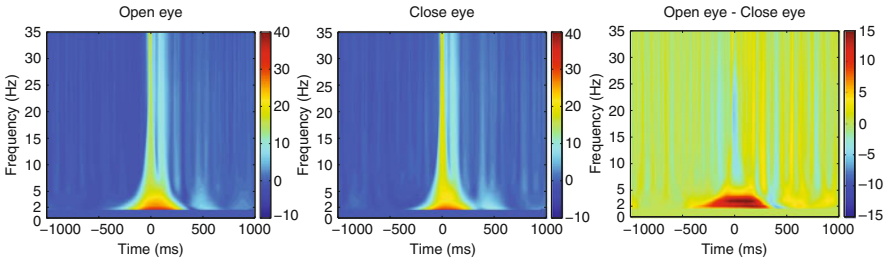


Fig. 1 Time-frequency diagrams of PLV_z for open eye, closed eye conditions and difference between them; average across all 19 ch from 10 participants. TMS shots were delivered at 0 ms, although TMS-evoked phase resetting precedes the TMS shots because of the limit of time resolution of the time frequency analysis

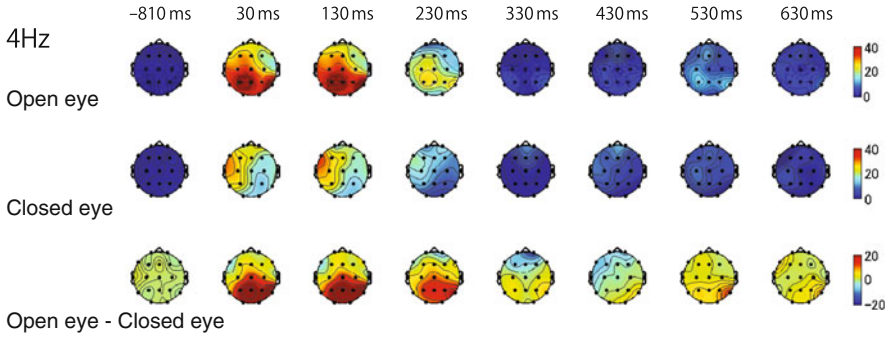


Fig. 2 Spherical-spline-interpolated PLV_z at various times before and after TMS at 4 Hz. PLV_z is significant ($p < 0.05$) if $\sigma > 3.01$ or < -3.01 (Bonferroni corrected) for open eye and closed eye. PLV_z is significant ($p < 0.05$) if $\sigma > 4.25$ or < -4.25 (Bonferroni corrected) for difference headmaps (*open–closed*)

This work was supported by MEXT Grant-in-Aid for Scientific Research (B) 193200207.

References

1. von der Malsburg, C.: The correlation theory of brain function. In Schulten K., ed.: Models of neural networks II. Berlin: Springer (1994) pp. 95–119.
2. Hirakura, Y., Yamaguchi, Y., Shimizu, H., Nagai, S.: Dynamic linking among neural oscillators leads to flexible pattern recognition with figure-ground separation. Neural. Netw. **9** (1996) 189–209.
3. Varela, F.J., Lachaux, J.P., Rodriguez, E., Martinerie, J.: The brainweb: Phase synchronization and large-scale integration. Nat. Rev. Neurosci. **2** (2001) 229–239.
4. Ward, L.M.: Synchronous neural oscillations and cognitive processes. Trends. Cogn. Sci. **7** (2003) 553–559.
5. Rodriguez, E., George, N., Lachaux, J.P., Martinerie, J., Renault, B., Varela, F.J.: Perception's shadow: Long-distance synchronization of human brain activity. Nature **397** (1999) 430–433.
6. Doesburg, S.M., Kitajo, K., Ward, L.M.: Increased gamma-band synchrony precedes switching of conscious perceptual objects in binocular rivalry. Neuroreport **16** (2005) 1139–1142.
7. Kitajo, K., Doesburg, S.M., Yamanaka, K., Nozaki, D., Ward, L.M., Yamamoto, Y.: Noise-induced large-scale phase synchronization of human brain activity associated with behavioural stochastic resonance. Europhys. Lett. **80** (2007) 40009–1–6.
8. Doesburg, S.M., Roggeveen, A.B., Kitajo, K., Ward, L.M.: Large-scale gamma-band phase synchronization and selective attention. Cereb. Cortex. **18** (2008) 386–396.
9. Mizuhara, H., Yamaguchi, Y.: Human cortical circuits for central executive function emerge by theta phase synchronization. NeuroImage **36** (2007) 232–244.
10. Massimini, M., Ferrarelli, F., Huber, R., Esser, S.K., Singh, H., Tononi, G.: Break down of cortical effective connectivity during sleep. Science **309** (2005) 2228–2232.
11. Tass, P., Rosenblum, M.G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., Freund, H.-J.: Detection of n:m phase locking from noisy data: Application to magnetoencephalography. Phys. Rev. Lett. **81** (1998) 3291–32.

Part III
Neural Coding and Realistic Neural Network Dynamics

Stimulation Induced Transitions in Spontaneous Firing Rates in Cultured Neuronal Networks also Occur in the Presence of Synaptic Plasticity Blocker KN93

Linda van der Heiden, Ildiko Vajda, Jaap van Pelt, and Arjen van Ooyen

Abstract Spontaneous firing activity in dissociated rat cortical tissue cultured in vitro shows highly stable firing rates over periods of hours. Recently, it was shown that a short period of low-frequency electrical stimulation induces significant and lasting changes in these firing rates. Now, it is shown that these changes also occur in the presence of the synaptic plasticity blocker KN93 in the culture medium. Apparently, the changes in firing rates after a short period of low-frequency stimulation do not depend on CaMK-II mediated synaptic plasticity.

Keywords Cultured networks · Spontaneous firing · Stable firing rates · Attractor dynamics · State transitions

1 Introduction

When dissociated rat cortical tissue is brought into culture, neurons readily grow out by forming axonal and dendritic arborizations and synaptic connections, and display spontaneous firing activity from about the end of the first week in vitro. This activity typically consists of alternating periods of synchronized (network bursts) and of largely uncorrelated activity. Mean firing rates at the individual sites of a multi-electrode array appear to be stable over periods of hours, changing only on developmental time scales, [1–4]. Recently, Vajda et al. [5] showed that a short period of low frequency (LF) electrical stimulation induces significant and lasting changes in these mean firing rates. The possible role of synaptic plasticity in these activity changes has now been investigated by repeating the experiments in [5] in the presence of the plasticity blocker KN93 in the culture medium. KN-93 prevents both pre- and postsynaptic CaMK-II activation, necessary for LTD and LTP induction, respectively [6–8]. The present experiments confirm the results of [5], also in the presence of KN93.

L. van der Heiden (✉)
Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_24,

2 Methods

Cell cultures on multielectrode arrays – Dissociated E18 rat neocortical neurons were cultured on planar multi-electrode arrays (MEAs) from MultiChannel Systems e.g., [1]. The preparation of the cultures was similar as in [5]. MEAs had either a so-called Hexagonal configuration with mixed 30, 20 and 10 μm electrode diameters or an 8 × 8 configuration with 10 μm micrometer diameter.

Pattern of electrical stimulation – The low-frequency (LF) electrical stimulation protocol, as shown and explained in Fig. 1, was derived from [5]. During the present experiments the culturing medium has been replaced by the recording medium Hepes ($N = 16$) or by Hepes + 10 μM KN93 ($N = 11$).

Comparison of mean firing rates at individual sites between two periods - Mean firing rates at individual sites in two periods were compared using a scatter plot as shown in Fig. 2 [5]. Data points at the diagonal indicate sites with equal firing rates for both periods. Dashed lines indicate 3 standard deviation (3σ) boundaries assuming a Poisson distributed spike train. The relative deviation *rd* of a data point from the diagonal is expressed by the ratio of its distance from the diagonal and the 3σ distance. The mean relative deviation *mrd*, of all data points in the scatterplot, thus provides a measure for the amount of scatter around the diagonal. The expected *mrd* value for a Poisson distributed spike train is equal to 0.27 [5]. For each experiment, the following periods of spontaneous activity were compared: (a) the 1st and 2nd half of the period before stimulation (Fig. 2, first column), b the 2nd half of the period before stimulation and the 1st half of the period after stimulation (Fig. 2, central column), and c the 1st and 2nd second half of the period after stimulation (Fig. 2, right column).

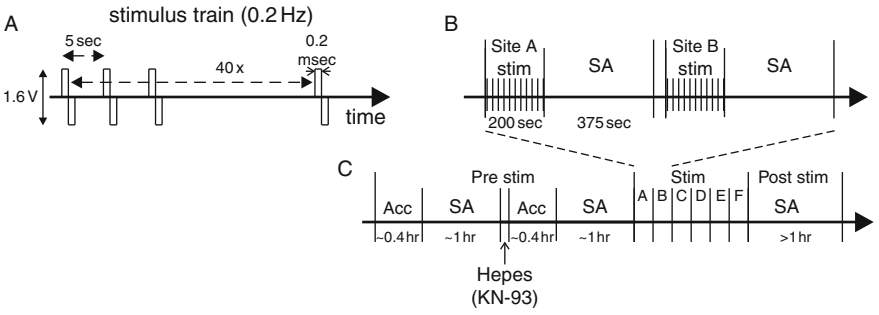


Fig. 1 **a** Pattern of low-frequency stimulation, consisting of a train of 40 bipolar pulses of 1.6 V (peak–peak) of 0.2 ms width (*single phase*), delivered with 5 s intervals (0.2 Hz). **b** Each pulse train is followed by a period of 375 s for recording spontaneous activity (SA) in the network. **c** Pulse trains are successively applied to six different electrodes in the multielectrode array. At the start of the experiment and after replacing culturing medium by recording medium Hepes (without or with KN93) a period of accommodation (Acc) is included. Both before and after the period of stimulation spontaneous activity is recorded for a period of at least 1 h

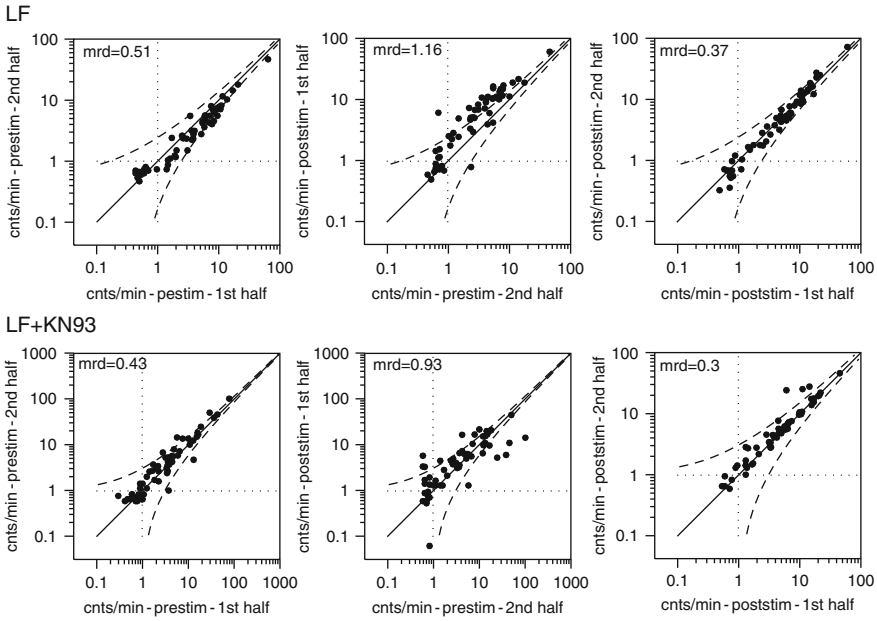


Fig. 2 Scatter plots comparing mean firing rates at individual sites between two periods. Compared are the 1st and 2nd half of the pre-stimulus (prestim) period (*left column*), the 2nd half of the prestim period with the 1st half of the post-stimulus (poststim) period (*central column*), and the 1st and 2nd half of the poststim period (*right column*). Data points at the diagonal indicate sites with equal firing rates in both periods. *Dashed lines* indicate 3σ (standard deviation) boundaries assuming a Poisson distributed spike train. The relative deviation *rd* of a data point from the diagonal is the ratio of its distance from the diagonal and the 3σ distance. The mean of all *rd* values in a scatterplot (*mrd*) is displayed in the upper *left* corner of each panel

3 Results

Scatterplots of low frequency (LF) stimulation experiments – Fig. 2 shows and describes examples of scatterplots for the three types of comparisons (columns) of firing rates (pre–pre, pre–post, post–post) in the absence (top row – LF) and in the presence of KN93 (bottom row - LF+KN93).

Analysis of mrd values. A summary of the *mrd* values is shown in Fig. 3, as well as the statistical comparisons between the different groups, as labeled by "a", "b", and "c", indicating the different periods in the firing rate comparisons (see scheme in Fig. 3). In both the LF and the LF+KN93 experiments, the "b" group of *mrd* data has significantly higher mean values than the "a" or "c" groups. This is not the case in the right panel of control experiments, in which group "b" compares firing rates in periods separated by a control period of the same duration as a stimulation period (data from experiments without or with KN93). The LF-results confirm the finding in [5], that low-frequency stimulation induces significant and lasting changes in the spontaneous firing rates in the network. Importantly, the LF+KN93 experiments

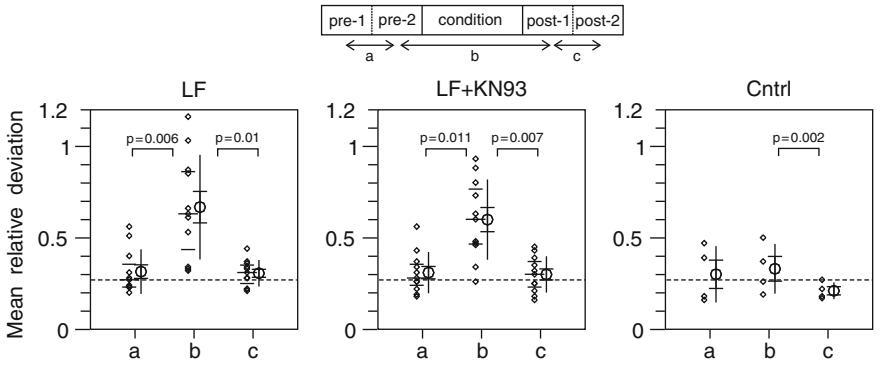


Fig. 3 Mean-relative-deviation (*mrd*) values in the absence (LF) and in the presence (LF+KN93) of the plasticity blocker KN93, and for the control experiments (Cntrl). Individual *mrd* values are shown along with their median and quartiles, as well as the mean, the SD intervals and the standard error in the mean intervals. The *p*-values are obtained from using the Kruskal Wallis test. The *dashed lines* indicate the *mrd* value of 0.27 as expected from fluctuations in firing rates in a Poisson distributed spike train

show that the plasticity blocker KN93 does not prevent these changes to occur. The dashed lines in Fig. 3 indicate the *mrd* value of 0.27 for the fluctuations in firing rates as expected for a Poisson distributed spike train. The *mrd* values of the test comparisons "a" and "c" appear to be close to this theoretical expectation, underscoring the stability in firing rates during a period of spontaneous activity.

4 Discussion

The present results confirm the findings of [5], that a period of low-frequency electrical stimulation induces significant and lasting changes in the pattern of spontaneous firing rates, and also demonstrate them in the presence of the plasticity blocker KN93. Apparently, these activity changes do not depend on CaMK-II mediated synaptic plasticity, suggesting that they are not caused by changes in synaptic efficacies. Other intrinsic network properties may have changed as well by low-frequency stimulation. However, the activity changes may also be of a dynamical nature, i.e., caused by state transitions in the firing dynamics state space, without structural changes in the neuronal network. Further research is needed to substantiate this interesting hypothesis.

References

1. Van Pelt, J., Wolters, P.S., Corner, M.A., Rutten, W.L.C., Ramakers, G.J.A.: Long- term characterization of firing dynamics of spontaneous bursts in cultured neural networks. IEEE-TBME **51** (2004) 2051–2062.

2. Segev, R., Baruchi, I., Hulata, E., Ben Jacob, E.: Hidden neuronal correlations in cultured networks. Phys. Rev. Lett. **92** (2004) 118102.
3. Chiappalone, M., Bove, M., Vato, A., Tedesco, M., Martinoia, S.: Dissociated cortical networks show spontaneously correlated activity patterns during in vitro development. Brain. Res. **1093** (2006) 41–53.
4. Le Feber, J., Van Pelt, J., Rutten,W.L.C.: Latency-related development of functional connections in cultured cortical networks. Biophys. J. **96** (2009) 3443–3450.
5. Vajda, I., Van Pelt, J.,Wolters, P., Chiappalone, M., Martinoia, S., van Someren, E., van Ooyen, A. Low-frequency stimulation induces stable transitions in stereotypical activity in cortical networks. Biophys. J. **94** (2008) 5028–5039.
6. Lisman, J., Schulman, H., Cline, H.: The molecular basis of CaMKII function in synaptic and behavioural memory. Nat. Rev. Neurosci. **3** (2002) 175–190.
7. Stanton, P.K., Gage, A.T.: Distinct synaptic loci of Ca2+/calmodulin-dependent protein kinase II necessary for long-term potentiation and depression. J. Neurophysiol. **76** (1996) 2097–2101.
8. Ninan, I., Arancio, O.: Presynaptic CaMKII is necessary for synaptic plasticity in cultured hippocampal neurons. Neuron **42** (2004) 129–141.

Attractors in Neurodynamical Systems

Włodzisław Duch and Krzysztof Dobosz

Abstract Analysis of the number, type and properties of attractors in complex neurodynamical systems is quite difficult. Fuzzy Symbolic Dynamics (FSD) creates visualization of trajectories that are easier to interpret than recurrence plots, showing basins of attractors. The variance of the trajectory within the attraction basin plotted against the variance of the synaptic noise provides some information about sizes and shapes of their basins. Semantic layer of dyslexia model implemented in the Emergent neural simulator is analyzed.

Keywords Neurodynamics · Attractor networks · Symbolic dynamics · Multidimensional time series visualization

1 Introduction

A general method for analysis of dynamical systems is based on recurrence plots [1]. A trajectory $\mathbf{x}_i = \mathbf{x}(t_i)$ returns to almost the same area (within ϵ distance) at some other point in time, creating non-zero elements (or black dots in the plot) of the recurrence matrix $\mathbf{R}_{ij}(\epsilon) = \Theta(\epsilon - ||\mathbf{x}_i - \mathbf{x}_j||)$. With some experience such plots allow for identification of many interesting behaviors of the system. However, these plots depend strongly on the choice of ϵ and the choice of the time step $\Delta t = t_{i+1} - t_i$. They may become irregular, the patterns in the plot may wash out, show spurious correlations or tangential motions [1], especially if steps along the trajectory $\mathbf{x}(t_i)$ are not constant for consecutive t_i.

Complex dynamics is frequently modeled using attractor networks, but precise characterization of attractor basins and possible transitions between them is rarely attempted. Another point of view on global analysis is quite fruitful. Recurrence plots use as the reference previous points $\mathbf{x}_i$ on the trajectory. This is analogous to

W. Duch (✉)
Department of Informatics, Nicolaus Copernicus University, Toruń, Poland
e-mail: wduch@is.umk.pl

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_25,

the use of the single-nearest neighbor method in classification [2]. Description of the density of the trajectories may be simplified and improved by selecting lower number of reference points and associating basis functions with these points. The recurrence rule $\mathbf{R}_i(\mathbf{x};\epsilon) = \Theta(\epsilon - ||\mathbf{x}_i - \mathbf{x}||)$ may then be reinterpreted as the return of the trajectory $\mathbf{x}(t)$ into the neighborhood of the vector $\mathbf{x}_i$ associated with some metric function $||\cdot||$ and some parameters ϵ. In particular if L_∞ (Chebyshev) metric is used the neighborhoods are hyperboxes and positive $\mathbf{R}_i(\mathbf{x};\epsilon)$ entry is marked by a symbol s_i, changing the vector trajectory into a sequence of symbols. This approach, known as the symbolic dynamics [3] (SD), has been used to simplify description of relatively simple dynamical systems, providing very rough approximation to the description of the trajectory. However, such discretization may be too rough for most applications. The Fuzzy Symbolic Dynamics (FSD) introduced by us recently [4] replaces hyperboxes by membership functions that estimate how far is the trajectory from the reference points. This is in fact a projection of the trajectory on a set of k basis functions $G(\mathbf{x};\mu_i,\sigma_i), i = 1..k$, localized around μ_i with some parameters σ_i, strategically placed in important points of the phase space.

FSD provides a non-linear reduction of dimensionality suitable for visualization of trajectories, showing many important properties of neurodynamics, such as the size and the relative position of attractors. To analyze more precisely properties of individual attractors response of the system to different type of noise with increasing variance is studied. A model of word recognition implemented in the Emergent simulator [5] is used for illustrative purpose. Conclusions about the relations between properties of attractors for different words are drawn.

2 Visualization of Attractors

For FSD visualization two Gaussian membership functions $G(\mathbf{x};\mu_i,\sigma_i), i = 1,2$ are placed in different parts of the phase space [4]. These functions are activated each time system trajectories pass near their centers, providing a sequence of fuzzy symbols along the trajectory.

As an example of what one can learn from such mapping visualization of a large semantic layer in the model of dyslexia implemented in the Emergent simulator is presented (see [5], chapter 10). This model has full bidirectional connectivity between orthography (6×8 units), phonology (14×14), and semantic layers (10×14), with recurrent self-connections within each of these layers, and additional hidden layers of neurons between each of these 3 layers. The model has been trained on 40 words, half of them concrete and half abstract. Semantics has been captured by using 67 features for concrete words (with average of 18 active features per word) and 31 for the abstract ones (about 5 active features on average), with half of the semantic layer devoted to abstract and half to concrete features. Correlation dendrogram between all 40 words is presented in [5], Chapter 10, Fig. 10.7. In the model all lesion parameters were off because the goal was to show attractors which have evolved in the learning process for different words (effects of lesions are reflected in visualizations, but are not shown here).

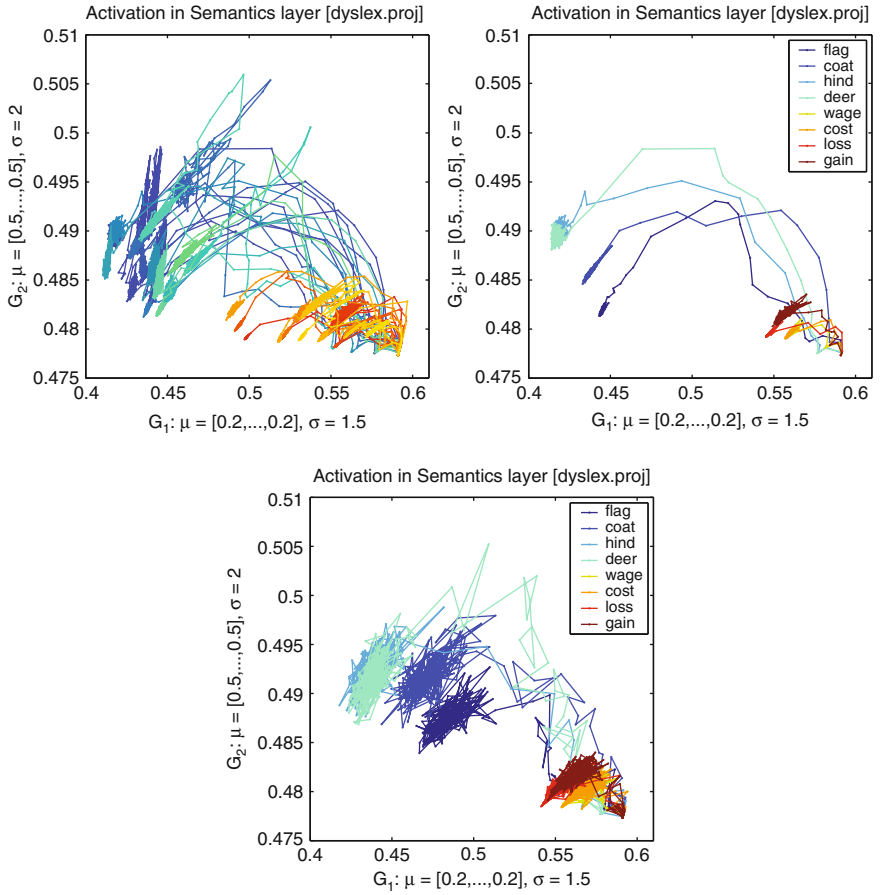


Fig. 1 FSD mapping showing trajectories and attractors of all 40 words used in the model (*left*); abstract words are plotted in the *right lower*, concrete words are in the *left-middle area*. Selected 8 words with the variance of the noise increased from 0.02 to 0.09

Two Gaussian functions with rather large dispersions have been placed in the 140 dimensional space of semantic layer activations. Parameters of these functions are given in Fig. 1, where trajectories for all 40 words are shown. The division between abstract and concrete words leads to quite different representations. Abstract words reach much faster their final configurations due to the significantly lower number of active units, while the meaning of concrete words is established only after a long trajectory, with more complex configurations.

Plots in Fig. 1 also show different shapes of attractor basins. The dyslexia model with default parameters (e.g. without any noise) is deterministic and converges to the point attractor, because the model has already been trained and starts from zero activity except for the input layer. To show attraction basins for a given word

Gaussian synaptic noise with zero mean and progressively growing variance has been introduced. In Fig. 1 (center, right) trajectories and attractor basins of 4 pairs of correlated words are displayed, with concrete words in the first two pairs (flag, coat, hind, deer) and abstract words (wage, coat, loss, gain) as the last two pairs. The variance of the noise has been increased from 0.02 to 0.09.

From such plots (due to the lack of space only a few are shown here) a number of interesting observations is made:

– convergence to sparse, simple representations (abstract words) is much faster than to more complex ones;
– some attractors may be difficult to reach – this is indicated by chaotic trajectories that lead to them;
– semantic representations from pairs of similar words are usually quite close to each other, with (coat, flag) being the most distinct pair;
– shapes of attraction basins differ for different words and noise levels – this is seen when trajectories explore larger areas with increasing noise levels;
– noise with a small variance explores only the bottom of attraction basin, for stronger noise the system may still show distinct attractors although patterns of semantic activity are severely distorted.
– for very strong synaptic noise (> 0.14) plots of all basins of attractors shrink.

The last observation has been quite surprising, but this is the result of quite complex dynamics of the model, with strong random saturation of the units and inhibition of most of them. Plot shown in Fig. 2 of the variance of trajectory near the attractor basin center as a function of increasing noise variance can serve as a measure of the basin depth. Initially there is little change in the trajectory variance,

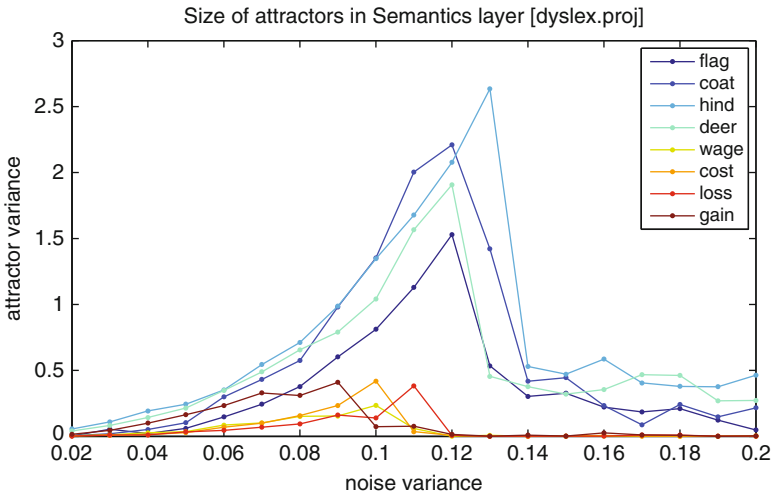


Fig. 2 Variance of the trajectory in the basin center as a function of increasing noise variance

showing that these basins are quite deep near the center, but for larger noise level there is sharp increase, as is also seen in Fig. 2, showing that the upper parts of the basin are much larger and additional energy helps to explore them. A sudden decrease of the trajectory variance is observed in all cases when synaptic noise is very strong.

3 Discussion

Cognitive science has reached the point where moving from finite automata models of behavior to continuous cognitive dynamics becomes essential [6]. Recurrent plots and symbolic dynamics may not be sufficient to show various important features of neurodynamics in complex systems. In this paper two new techniques have been explored: first, a fuzzy version of symbolic dynamics, that may also be treated as smoothed version of recurrent plot technique with large neighborhood, and second, analysis of variance of trajectories around point attractor as a function of noise. FSD depends on the choice of membership functions, and thus has more parameters than recurrent plots or the discrete symbolic dynamics [3] that may be obtained from the fuzzy version by thresholding the activations of membership functions. However, FSD retains more information about dynamics and creates plots that are easier to interpret than recurrence plots, less dependent on the sampling step.

For quasi-periodic attractors variance in the direction perpendicular to the trajectory may be estimated. Perhaps for the first time these techniques may show how attractor basins created by various learning procedures depend on the similarity of stimuli, their context, but also on the properties of neurons. For example, reducing the leaky ion channel currents creates basins attracting trajectories strongly and preventing attention shifts, a phenomenon that is of fundamental importance in autism.

References

1. Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. Phys. Rep. **438** (2007) 237–329.
2. Shakhnarovish, G., Darrell, T., Indyk, P., eds.: Nearest-Neighbor Methods in Learning and Vision. Cambridge, MA: MIT Press (2005).
3. Hao, B., Zheng, W., eds.: Applied Symbolic Dynamics and Chaos. Singapore: World Scientific (1998).
4. Dobosz, K., Duch,W.: Understanding Neurodynamical Systems Via Fuzzy Symbolic Dynamics. Neural. Netw. **23** (2010) 487–496.
5. O'Reilly, R.C., Munakata, Y.: Computational Explorations in Cognitive Neuroscience. Cambridge, MA: MIT Press (2000).
6. Spivey, M.: The Continuity of Mind. New York, NY: Oxford University Press (2007).

Combining Supervised, Unsupervised, and Reinforcement Learning in a Network of Spiking Neurons

Sebastian Handrich, Andreas Herzog, Andreas Wolf, and Christoph S. Herrmann

Abstract The human brain constantly learns via mutiple different learning strategies. It can learn by simply having stimuli being presented to its sensory organs which is considered unsupervised learning. In addition, it can learn associations between inputs and outputs when a teacher provides the output which is considered as supervised learning. Most importantly, it can learn very efficiently if correct behaviour is followed by reward and/or incorrect behaviour is followed by punishment which is considered reinforcement learning. So far, most artificial neural architectures implement only one of the three learning mechanisms — even though the brain integrates all three. Here, we have implemented unsupervised, supervised, and reinforcement learning within a network of spiking neurons. In order to achieve this ambitious goal, the existing learning rule called spike-timing-dependent plasticity had to be extended such that it is modulated by the reward signal dopamine.

1 Introduction

The human brain is able to establish memory representation on the basis of dynamical changes of synaptic connections between neurons. Such synaptic memory representations can be built up by multiple ways of learning. On the one hand, it is possible that repeated presentation of stimuli leads to stronger synaptic connections. Hebb already postulated this form of learning in 1949 and suggested a learning rule which said that synaptic connections between two neurons increase, when both neurons are firing at the same time (neurons that fire together, wire together) [1]. This form of learning requires no feedback from the environment telling the organism whether the changed synaptic connection leads to a behavioural advantage or a reward (pain minimization or pleasure maximization). Therefore, this form of

S. Handrich (✉)
Department for Biological Psychology, Institute for Psychology II, Otto-von-Guericke-University Magdeburg, 39106 Magdeburg, Germany
e-mail: sebastian.handrich@ovgu.de

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_26,

learning is referred to as "unsupervised learning" in the domain of artificial neural networks. On the other hand, there is a form of learning requiring the organism to learn a specific task, e.g. a categorization that was previously unknown. Here, the correct classification of each stimulus must be provided. Thus, this form of learning is referred to as "supervised learning". In Psychology, implicit memory with effects of repetition priming result from this type of learning. The last and most frequent form of learning is "reinforcement learning". Here, feedback from the environment is given that can either be appetitive (reward) or aversive (punishment). Classical conditioning belongs to this class of learning [2]. In supervised learning, the feedback provides the correct response (e.g. category) that the learner (e.g. network) should give. In reinforcement learning, the feedback merely indicates the correctness (right or wrong via reward or punishment) but not the correct response.

The human brain integrates frequently and without effort all three learning strategies. However, most neural implementations focus on only one out of the three, since each strategy requires specific features to be present in the neural architecture. The aim in the present paper is to create a network architecture which can learn via supervised, unsupervised, and reinforcement learning.

For our simulations, we will use biologically plausible, spiking neurons based on the model of Izhikevich [3]. The basic learning mechanisms at the synapse level will be spike timing dependent plasticity (STDP) [4, 5]. STDP is a refinement of Hebb's learning rule. Only if a pre-synaptic action potential leads to a post-synaptic potential, a causal relationship can be assumed and the synapse is strengthened. If, however, the post-synaptic potential precedes the pre-synaptic action potential this must be considered as acausal and the synapse will be weakened. It has been demonstrated that STDP can be used in networks of spiking neurons to implement unsupervised learning [6].

In an architecture that is able to learn via unsupervised learning, it should be easy to implement also supervised learning by simply activating the correct neuron of the output layer at the same time as the input arrives in the input layer. However, the temporal delay between pre-synaptic and post-synaptic activity must be positive in order to result in a strengthening of the synapse. Thus, it is more advantageous to inhibit the wrong output neurons. As a consequence it is assured that the correct neurons only fire due to preceding activity of input neurons.

Reinforcement learning requires three prerequisites within the neural architecture which are described in the following subsections. The first two have been solved by us recently and we now focus on the third.

1.1 Determining the Reward or Punishment

If the learning rule is supposed to depend upon the correctness of the network's output, it should be possible to unambiguously translate the output into a reward or punishment signal. However, while winner-take-all (WTA) architectures can achieve this goal in artificial neural networks [7], for spiking neurons WTA

implementations are not trivial. We have therefore developed a biologically plausible WTA algorithm (bpWTA) [8]. This bpWTA algorithm assures that only one neuron within the output layer of a spiking neural network fires while the others remain silent. For a classification problem the correctness of the network output can easily be determined and as a result decided whether to apply a reward or punishment signal to the network.

1.2 Solving the Distal Reward Problem

A further problem of reinforcement learning lies in the fact that the output might become active only after the input neurons have already stopped to fire. The temporal gap between input and output is especially relevant if hidden layers exist in the network. Therefore, the STDP rule can no longer determine the temporal relationship between pre- and post-synaptic activity. This temporal discrepancy has been termed the problem of delayed or distal reward [2, 9]. The problem has previously been solved by inserting a variable in a network of spiking neurons that stores the computed change of synaptic weights and decays over time [10]. However, we have developed a different approach [11]. We implemented a feedback layer that keeps the input layer active after the input has stopped. Our solution is well in line with electrophysiological data. Neurons in the inferior temporal cortex fire for a few hundred milliseconds after the end of external stimulation [12] and neurons in entorhinal cortex even longer [13]. However, in the current network, we did not need a feedback layer, since we had have no hidden layer and the input was active long enough for the output to be evaluated.

1.3 Switching from LTP to LTD Depending on Reward or Punishment

If an output neuron fires resulting in a correct solution or classification, a reward signal should be generated and the synapses from those neurons of preceding layers that led to the firing of the output neuron should be strengthened, i.e. long-term potentiation (LTP) should have taken place. If, however, an output neuron fires resulting in an erroneous solution or classification, a punishment signal should be generated and the synapses from those neurons that led to the firing of this output neurons should be weakened, i.e. long-term depression (LTD) should take place. When using the STDP learning rule, LTP would need to be inverted to LTD in case of punishment. This seems physiologically unplausible at first glance. However, electrophysiologists were recently able to demonstrate that dopamine, which plays a key role in reward signaling [14], also modulates STDP [15]. This finding was subsequently transformed into a revised STDP learning rule, called spike-timing-dependent eligibility (STDE) [16]. We have applied this new STDE and have used it in combination with the above-mentioned bpWTA to implement reinforcement learning in a network of spiking neurons.

2 Methods

2.1 Network Configuration

Our network architecture consists of two different layers, an input- and a WTA-layer. The input layer is used to receive the external stimuli as described below. It is composed of 64 excitatory neurons, that are not laterally connected. The WTA layer consists of two excitatory neurons, which (a) integrate their inputs received from the input layer and (b) inhibit each other via two inhibitory neurons. This leads to a binary categorization of the presented input patterns and allows to compute an error signal, needed for reinforcement learning. Each input neuron i is connected to each excitatory WTA neuron j with initial synaptic weights $c_{ij} = r^2/64$, where r is randomly chosen from the uniform distribution within the interval of $[0, 0.5]$. In the WTA layer all synaptic weights are set to $c_{WTA} = c_{MAX} = 1$.

There are two types of input to our network. The first one is the stimulation of the input layer defined by a grid of 8-by-8 sensitive areas that can be stimulated by the input patterns. Those patterns, if considered as images, show randomly distributed noise superimposed onto vertical or horizontal bars. Each sensitive area excites the corresponding input neuron by a Poisson distributed spike train (firing rate $\lambda_{ext} = 25$ Hz). The two patterns were presented alternately for one second with an inter-stimulus-interval (ISI) of one second. Additionally, each neuron receives thalamic background activity, generated by a Poisson-point process with a mean firing rate of $\lambda_{int} = 0.1$ Hz.

2.2 Neuron Model

We used the neuron model by Izhikevich [3], which reduces the biophysically accurate Hodgkin-Huxley model to a two-dimensional system of ordinary differential equations. The advantages of this model are the fast computation and the configuration of different behaviors via a few parameters. So, it is widely used e.g. to apply simulations of network development [17].

The neurons of the input layer and the inhibitory ones of the WTA-layer were modelled by:

$$\begin{aligned} \dot{v} &= 0.04v^2 + 5v + 140 - u - I_{syn}, \\ \dot{u} &= a\,(bv - u) \end{aligned} \tag{1}$$

where v is the membrane potential, u the recovery variable and I_{syn} is the total synaptic current. For further details see [8].

2.3 Neuron Types

In our study, we used three different types of neurons:

1. *Glutamate (excitatory) neurons* with, according to the simulations in [6]: $(a, b) = (0.02, 0.2)$ and $(c, d) = (-65, 8) + (15, -6)\,r^2$, where r is randomly

selected from the uniform distribution in the interval $[0, 0.5]$ to obtain a behavior between regular spiking (RS, $r = 0$) and intrinsically bursting (IB, $r = 0.5$). The square of r biases the distribution towards the RS cells.
2. *GABAergic (inhibitory) neurons*:
   $(a, b) = (0.02, 0.25)$ and $(c, d) = (-65, 2)$, to get low-threshold spiking (LTS) neurons.
3. *Input-integrating (excitatory) neurons*:
   $(a, b) = (0.02, -0.1)$ and $(c, d) = (-55, 6)$, used in the WTA-layer.

2.4 Synapse Model

The synaptic input current in a neuron is calculated in each time step by:

$$\begin{aligned} I_{syn} = & \, g_{AMPA}\,(v - 0) \\ & + g_{NMDA}\,\frac{[(v+80)/60]^2}{1+[(v+80)/60]^2}\,(v - 0) \\ & + g_{GABA_A}\,(v + 70) \\ & + g_{GABA_B}\,(v + 90)\,, \end{aligned} \tag{2}$$

where g_k is the time dependent synaptic conductance and v the actual membrane potential. The conductances change by first-order linear kinetics

$$\dot{g}_k = -\frac{g_k}{\tau_k} \tag{3}$$

with time constants $\tau_k = 5$, 150, 6 and 150 ms for the simulated *AMPA*, *NMDA*, $GABA_A$ and $GABA_B$ receptors, respectively [6]. The rise time of currents is typically short and neglected.

If a spike is transmitted from presynaptic neuron i to postsynaptic neuron j, after a delay-time δ_{syn}, the conductances are updated depending on the type of presynaptic neuron, the synaptic efficiency $R_i w_i$ and the synaptic weight c_{ij}:

$$g_k \leftarrow g_k + c_{ij}, \tag{4}$$

The transmission-delay δ_{syn} results from the Euclidean distance between the neurons with an additional latency of 0.5 ms.

2.5 Synaptic Plasticity

There are two types of plasticity in our model: short-term and long-term. The former is achieved by modelling the synaptic efficiency $R_i w_i$, which describes the fractional amount of available neurotransmitter at each synapse. The used model was developed by Markram [4]:

$$R_{i,n+1} = R_{i,n}(1 - w_{i,n+1}) \exp\left(\frac{-\Delta t}{\tau_{rec}}\right) + 1 - \exp\left(\frac{-\Delta t}{\tau_{rec}}\right), \tag{5}$$

$$w_{i,n+1} = w_{i,n} \exp\left(\frac{-\Delta t}{\tau_{fac}}\right) + U\left(1 - w_{i,n} \exp\left(\frac{-\Delta t}{\tau_{fac}}\right)\right) \tag{6}$$

with $\tau_{rec} = 400$ ms and $\tau_{fac} = 1,200$ ms.

To enable learning in our model, which is considered as a change of the synaptic weights, we used a well known form of hebbian learning, the STDP [5]. Therefore we implement a model by Redgrave and Gurney [14], which assumes based on experimental data [15], that STDP is not only describable by the temporal order of presynaptic and postsynaptic spikes, but also that the level of dopamine and the corresponding receptor type (D1,D2) have to be taken into account. Gurney proposes the use of two different STDP-kernels (no dopamine and high dopamine), which are superimposed linearly in dependence of the actual level of dopamine (see Fig. 1). We combined this approach with the STDP-model by Morrison et al. [18] , who used *STDP traces*, so that not only the *nearest-neighbour*, but every preceding spike influences the change of the synaptic weight.

The use of plastic synapses may end in an unconstrained feedback-cycle due to the fact, that correlated pre- and postsynaptic firing leads to LTP which in turn increases that correlation [19]. In biological neural networks this stability problem

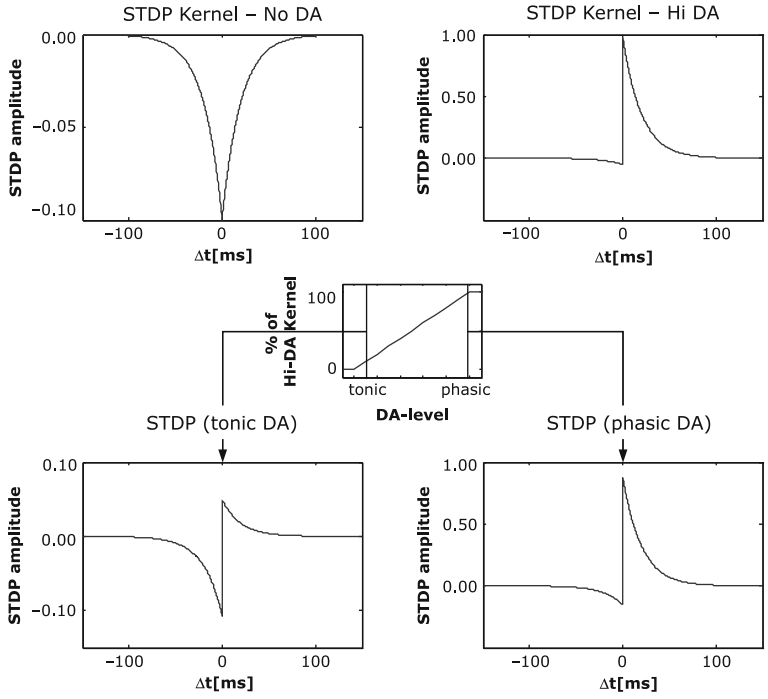


Fig. 1 Dopamine modulated STDP. The *upper pictures* show the two STDP kernel templates. The higher the level of dopamine, the higher the percentage of the high-dopamine-template on the resulting STDP function (*lower pictures*)

is avoided by a mechanism called *homeostatic-synaptic-scaling*, which changes the number of AMPA and NMDA receptors of the postsynaptic neuron in dependence of its activity [20]. This leads to a competitive scaling of the afferent synapses. In our model we achieved this competitive scaling by preventing the cumulative weights $c_\kappa = \sum_j^n c_{ij}$ of the synapses leading onto the same neuron to exceed a maximum weight of $c_{max} = 1$:

$$c_{ij} \leftarrow \frac{c_{ij}}{c_\kappa}, if\, c_\kappa \geq c_{max}. \tag{7}$$

In a further implementation, synaptic-scaling will be implemented as a neuron excitability adapted to its activity (firing rate).

2.6 Experiments

To apply three different learning mechanisms, we carried out three different experiments, i.e. simulations of our network architecture:

1. Unsupervised learning: We used a classification task, i.e. the network received no information about the category of the presented stimulus pattern. We expected, that similar inputs are mapped to the same excitatory output neuron, whereas different stimuli will excite different output neurons.
2. Supervised learning: The network had to learn to assign the presented stimulus pattern to a predefinded category. We considered two different ways to specify the correct category to the network: One can either excite the output neuron representing the desired category or inhibit all neurons representing the incorrect one. As mentioned above, we decided to chose the second way to avoid timing problems. We expected that the network learns to classify the stimuli correctly, because the STDP will strengthen only those synapses, that connect the input neurons, that were excited by the stimulus, and the desired, i.e. not inhibited, output neurons.
3. Reinforcement learning: In the case of a reward or punishment, the level of dopamine is set to $d_{High} = 0.6$ or $d_{None} = 0.0$, respectively. Then, the dopamine level decays to an equilibrium level $d_{equ} = 0.3$. The higher or lower level of dopamine results in a higher synaptic potentiation (LTP) or depression (LTD), respectively. If a stimulus is classified correctly or not, depends on which output neuron fires first.

In all three simulations we used exactly the same network architecture and the same initial synaptic weights as described above. There were two different basic stimulus patterns (vertical and horizontal bar). In every trial one of the basic patterns, randomly modified by 20% was presented to the network for 1 s. The inter-stimulus-interval (ISI) was also one second. There were 1,000 trials per simulation, so every experiment took 2,000 s (simulation time).

3 Results

The goal of our work was to develop a network architecture, that is able to perform three different learning mechanisms: unsupervised, supervised and reinforcement learning. First, we show the results of each learning method and then compare them with each other.

3.1 Unsupervised Learning

In Fig. 2, the result of the unsupervised learning is shown. One can see the membrane voltage of both the two excitatory output neurons $v_A(t)$, $v_B(t)$ and the two inhibitory WTA neurons $v_{AI}(t)$, $v_{AI}(t)$, driven by the excitatory ones. Also, the stimuli presented in each trial are depicted.

Due to the different weights and different stimulation, there are different input currents to the two excitatory WTA neurons, A and B. The excitatory WTA neuron with the highest input fires first, the related inhibitory neuron is activated with a short transmission delay and inhibits the opposite excitatory neuron. In an earlier work, we could show, that if the input contrast exceeds a threshold $C_{th} = 1.015$, in each trial only one of the excitatory WTA neurons is activated (see [8] for details).

The weights of the synaptic connections c_{ij} are initialized randomly, but normalized, so that no output neuron is initially prefered ($\sum_j c_{A,j} = \sum_j c_{B,j}$). Thus, at the beginning of the learning phase (trial 0–4) the contrast of the input received by the output neurons is very low, so only a small difference may determine which neuron is the winning one. This may lead to a wrong classification as depicted in Fig. 2: Although the stimulus in trial 3 was similiar to the one in trial 1, the wrong output neuron had won the WTA competition. However, as we could show in [8], even if

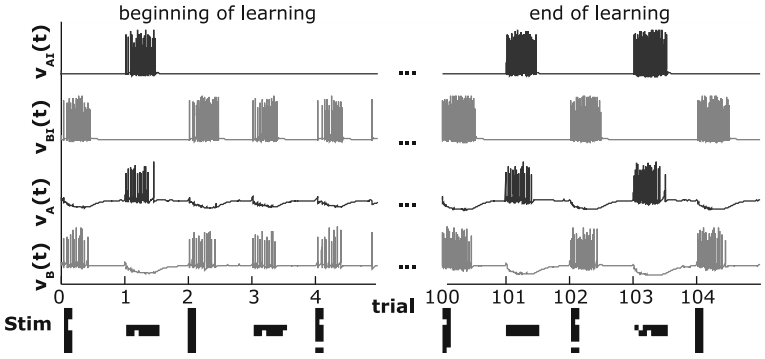


Fig. 2 Time course of the membrane potential of the excitatory neurons $v_A(t)$, $v_B(t)$ and inhibitory neurons $v_{AI}(t)$, $v_{BI}(t)$ of the WTA output layer during unsupervised learning. In the beginning of the learning phase (*left part*), there is no clear assignment between the presented stimulus and the activated output neuron, i.e. horizonzal and vertical input patterns both activate neuron B (trial 2 and 3). At the end of the learning phase (*right part*) this assignment is achieved

there is a low input contrast, there is a preference to the correct output neuron. Thus, if STDP is applied, the synapses between the input neurons, excited by the stimulus, and the correct output neuron are strengthened. This in turn leads to a higher input contrast, so a wrong classification becomes more unlikely. At the end of the learning phase (trial 100–104) always the neuron associated with the presented stimulus fires with a firing rate between 40 and 50 Hz.

3.2 Supervised Learning

To implement supervised learning we initially considered two different ways. First, one can excite the output neuron, representing the desired category. Thus, there is a simultaneous activation of the input and the desired output neuron, so that STDP will potentiate the connections between them. However, there is an error in this consideration: Since an input neuron does not fire only once, but with a poisson rate of $\lambda_i = 25$ Hz, an output spike may precede the input neuron spike, which leads to LTD and the corresponding synapse gets weakened. A solution of this problem is to excite the desired output neuron not with a fixed frequency, but time-delayed to the firing of the input neuron, but this is not biologically plausible. To circumvent this problem, we chose a second way: Instead of exciting the desired output neuron, we inhibited the incorrect ones.

In Fig. 3 the synaptic weights of the connections between the input layer and the two output neurons A and B are shown. At the beginning of the supervised learning (trial 0), the weights are randomly initialized. During the learning (trial 150), STDP potentiates only those synapses that connect the stimulus-driven input

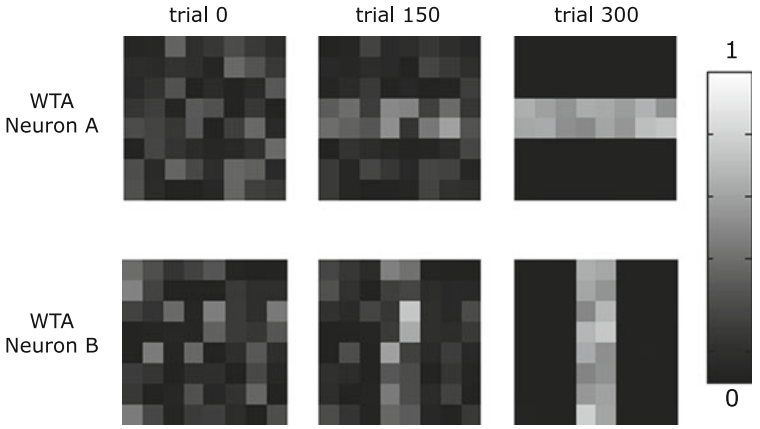


Fig. 3 Weight matrices of the connections from input- to output layer. *Left*: Random initialization. *Middle*: During supervised learning. The synapses representing the stimuli are potentiated. *Right*: End of learning phase. Weights of those synapses used for the stimuli representations remain constant on a high level. Unused synaptic connections are depressed

neurons and the firing (not inhibited) output neuron. At the end of the supervised learning (trial 300), the strength of the synapses that represent the stimuli is high. Due to the competitive synaptic scaling (see methods), all synapses, that are not used for a representation of the stimulis are weakened.

3.3 Reinforcement Learning

During the reinforcement learning, the network receives information, if it has categorized the presented stimulus correctly through a reward- or punishment signal, respectively. The decision, whether the categorization is correct is determined by the first spike of an output neuron in each trial. If the desired output neuron fires first, the network output is rewarded or punished, conversely. Here, the following problem occurs: Since the correctness of the output is not known until the first output neuron fires, a subsequent reward or punishment signal can not influence the synaptic weight, because the STDP has already applied. But there is no need to do this. Because the output neuron do not fire only once, but with a frequency f_{fire} (40–50 Hz at the end of the learning phase), the reward modulated STDP will influence the subsequent input-output spike pairings. Figure 4 shows the different network behavior at the beginning (trial 0–4) and at the end of the learning phase (trial 200–204). If the correct output neuron fires first in the respective trial, i.e. a horizontal pattern activates neuron A or a vertical pattern activates neuron B, we set the level of dopamine (*DA-level* in figure) to $d_{High} = 0.6$. This leads to a strong potentiation (LTP) of the synaptic connections between the stimulus-driven input neurons and the active output neuron. In trial 1 of the example shown in Fig. 4, the incorrect

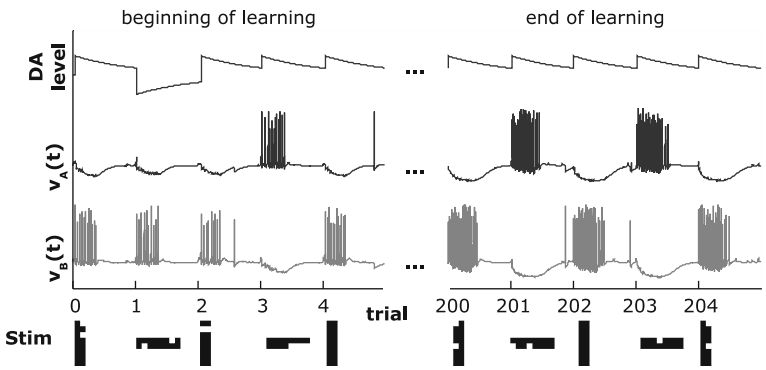


Fig. 4 Reinforcement learning: Time course of the membrane potential of the excitatory output neurons $v_A(t)$, $v_B(t)$ and the level of dopamine (DA-level) are shown. Correct, i.e. horizontal patterns activate neuron A, or incorrect, i.e.vertical patterns activate neuron B, categorization leads to an increased (reward) or decreased (punishment) level of dopamine, respectively. At the beginning of the learning phase (*left part*) wrong categorizations may occur. When the training is completed (*right part*), presented stimuli are categorized correctly and the firing frequency of the corresponding desired output neuron has increased to 40–50 Hz

output neuron fired, so the level of dopamine is decreased to $d_{None} = 0.0$. This leads to a synaptic depression (LTD). After applying the reward or punishment signal the level of dopamine decays to an equilibrium level, that produces an standard STDP course.

3.4 Comparison

To compare the three learning mechanisms with each other we considered the firing rate of the desired and the incorrect output neuron in each trial according to the presented stimulus (see Fig. 5). Due to the small synaptic weights the output neurons initially fire with a low frequency (about 3 Hz). During the learning phase the firing rate of the respective correct output neuron increases until it remains constant in the range between 40 Hz and 50 Hz (gamma band). However, the firing frequency does not converge with an equal rate in each learning method. In the case of unsupervised learning the firing frequency remains constant after 250 presentations of stimulus A or B, respectivly. When supervised learning is applied, it took 150 trials and during reinforcement learning the final output frequency is already achieved after only 50 stimulus presentations. This is an expected result since (a) during supervised learning the incorrect synapses are not as often potentiated by

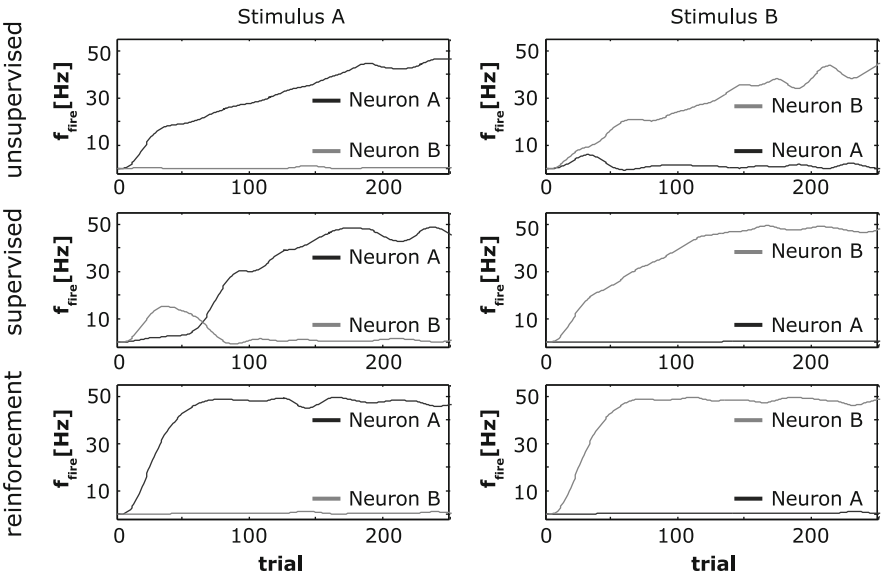


Fig. 5 Comparison of the three learning mechanisms. The firing frequencies of the correct and incorrect output neuron respective to the presented stimuli are shown. In the case of reinforcement learning the final fire rate is archieved after only 50 stimulus presentations (*bottom row*). Supervised learning works slower (150 trials, *middle row*) and in the case of unsupervised learning it took more than 250 trials to archieve the final state (*top row*). In every experiment the final firing rate was in the range of 40–50 Hz

STDP as during unsupervised and (b) during reinforcement learning the change of the synaptic weights is, much higher compared to unsupervised learning due to the dopamine modulated STDP.

4 Discussion and Conclusion

We have implemented a spiking network architecture with a biologically plausible neuron model and conductance based synaptic connections which is able to perform different types of learning.

The output layer represents a bpWTA architecture. It processes the data from the input layer and assigns the different input patters to result categories. Depending on the correct categorization, dopamine acts as a reward signal to adjust the STDP rule. This enables us to adjust the weights of the connections between input layer and output layer according to three different types of learning (supervised, unsupervised, and reinforcement).

Many other models have used the mean firing rate of a population or of single cells as input to other cells, e.g. so-called mean field models [21–23]. Such models would have to determine the category via the neuron or population with the highest mean firing rate. However, for time intervals in the order of the inter-spike interval a mean firing rate can not be defined. We have designed a biologically plausible model with spiking neurons for two reasons. First, we wanted to apply a learning rule at the synapse level which is only possible when single neurons are modelled instead of mean fields. Second, we plan to compare the network activity with biological EEG experiments in future studies. Interestingly, this has also been achieved for networks of single neurons [24,25]. The selected integrator characteristic of our output neurons is biologically plausible and combines the mean firing rate of earlier spikes with a fast computation of present spikes. Thus, the computed category of an input pattern is represented by the excitatory output neuron which receives the most input from many input neurons within a specific integration time. The integration time depends on synaptic components (NMDA time constant) and parameters of the integrator neuron.

As we had demonstrated previously, the bpWTA architecture in combination with the STDP learning rule provides the possibility for unsupervised learning [8].

For supervised learning, the categories of the training examples are known and will be used as a teaching signal. We implemented the teaching signal by an inhibition of the wrong output neuron. This inhibition prevents the wrong output neuron from firing. Hence, only the correct output neuron has a chance to become active by the integration of its input and only the synapses of the firing neuron can undergo dynamical changes via STDP.

In order to implement also reinforcement learning within the same architecture, a reward signal has to be generated. Dopamine is commonly considered as a reward signal [26] and experiments suggest that a reduction of dopamine secretion may also serve as a punishment signal, e.g. [27]. Recent investigations [15] show, that the STDP curves depend on the level of dopamine. This behavior can be modelled as

different STDP kernels [16]. In the network model presented here, these dopamine sensitive STDP kernels successfully implement reinforcement learning.

Acknowledgments This work was supported by the DFG HE 3353/6-1 Forschungsverbund "Electrophysiological Correlates of Memory and their Generation", SFB 779, and by BMBF Bernstein-group "Components of cognition: small networks to flexible rules".

References

1. Hebb, D.: The Organization of Behavior: A Neuropsychological Theory. New York, NY: Wiley (1949).
2. Dayan, P., Abbott, L.F.: Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. Cambridge, MA: MIT press (2001).
3. Izhikevich, E.M.: Simple model of spiking neurons. IEEE. Trans. Neural. Netw. **14**(6) (2003) 1569–1572.
4. Markram, H., Lubke, J., Frotscher, M., Sakmann, B.: Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science **275**(5297) (1997) 213–215.
5. Bi, G.Q., Poo, M.M.: Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neurosci. **18**(24) (1998) 10464–10472.
6. Izhikevich, E.M., Gally, J.A., Edelman, G.M.: Spike-timing dynamics of neuronal groups. Cerebral. Cortex. **14** (2004) 933–944.
7. Yuille, A.L., Geiger, D.: Winner-Take-All Networks. In Arbib, M.A., ed.: The Handbook of Brain Therory and Neural Networks, 2nd edn. Cambridge, MA: MIT Press (2003) 1228–1231.
8. Handrich, S., Herzog, A., Wolf, A., Herrmann, C.S.: A Biologically Plausible Winner-Takesall Architecture. In: International Conference on Intelligent Computing (ICIC). (in press).
9. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press (1998).
10. Izhikevich, E.M.: Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb. Cortex. **17**(10) (2007) 2443–2452.
11. Handrich, S., Herzog, A., Wolf, A., Herrmann, C.S.: Prerequisites for integrating unsupervised and reinforcement learning in a single network of spiking neurons. In: International Joint Conference on Neural Networks (IJCNN). (in press).
12. Desimone, R.: Neural mechanisms for visual memory and their role in attention. Proc. Natl. Acad. Sci. USA. **93**(24) (1996) 13494–13499.
13. Suzuki,W.A., Miller, E.K., Desimone, R.: Object and place memory in the macaque entorhinal cortex. J. Neurophysiol. **78**(2) (1997) 1062–1081.
14. Redgrave, P., Gurney, K.: The short-latency dopamine signal: a role in discovering novel actions? Nat. Rev. Neurosci. **7**(12) (2006) 967–975.
15. Shen, W., Flajolet, M., Greengard, P., Surmeier, D.J.: Dichotomous dopaminergic control of striatal synaptic plasticity. Science **321**(5890) (2008) 848–851.
16. Gurney, K., Redgrave, P.: A model of sensory reinforced corticostriatal plasticity in the anaethetised rat. In: Society for Neuroscience Abstracts. (2008).
17. Herzog, A., Kube, K., Michaelis, B., de Lima, A., Baltz, Tand Voigt, T.: Contribution of the GABA shift to the transition from structural initializationto working stage in biologically realistic networks. Neurocomputing (2008) 1134–1142.
18. Morrison, A., Diesmann, M., Gerstner, W.: Phenomenological models of synaptic plasticity based on spike timing. Biol. Cybern. **98**(6) (2008) 459–478.
19. Abbott, L., Nelson, S.: Synaptic plasticity: taming the beast. Nat. Neurosci. **3**(Suppl) (2000) 1178–1183.

20. Turrigiano, G.: The self-tuning neuron: synaptic scaling of excitatory synapses. Cell **135**(3) (2008) 422–435.
21. Wright, J.J., Liley, D.T.: Simulation of electrocortical waves. Biol. Cybern. **72**(4) (1995) 347–356.
22. David, O., Friston, K.J.: A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage **20**(3) (2003) 1743–1755.
23. Freeman,W.J.: A field-theoretic approach to understanding scale-free neocortical dynamics. Biol.Cybern. **92**(6) (2005) 350–359.
24. beim Graben, P., Kurths, J.: Simulation of global properties of electroencephalograms with minimal random neural networks. Neurocomputing **71** (2008) 999–1007.
25. Fründ, I., Herrmann, C.S.: Simulating evoked gamma oscillations of human EEG in a network of spiking neurons reveals an early mechanism of memory matching. In Descalzi, O., Rosso, O., Larrondo, H., eds.: Nonequilibrium Statistical Mechanics and Nonlinear Physics, AIP Conference Proceedings 913(1), American Institute of Physics (2007) 215–221.
26. Schultz, W.: Dopamine neurons and their role in reward mechanisms. Curr. Opin. Neurobiol. **7**(2) (1997) 191–197.
27. Ungless, M., Magill, P., Bolam, J.: Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science **303**(5666) (2004) 2040–2042.

Concerted Activities in Frog Retinal Ganglion Cells

Wei Jing, Hai-Qing Gong, and Pei-Ji Liang

Abstract In the present study, the spike discharges of frog retinal ganglion cells in response to repetitive full field white light flashes were recorded using multi-electrode array. Cross-correlation analysis showed that two different types of concerted activities could be distinguished according to the peak width of the cross-correlation function. Further analysis revealed that the concerted activity pattern of ganglion cells was dependent on the distance between the neurons being examined: nearby cells often fired in precise synchrony, while correlated activities with distributed time lags were mainly observed from remote cell pairs.

Keywords Multi-electrode recording · Population activity · Cross-correlation

1 Introduction

All of our visual experience is dependent on the spike activities of retinal ganglion cells (RGCs). However, the number of RGCs is the fewest among all kinds of neurons in the visual pathway [1]. This anatomy limitation makes it necessary that the visual information is efficiently organized in the RGC layer before it is transferred to the central visual part. Concerted population activity is one of the important ways for the retina to organize visual information, it was suggested that concerted activity of RGCs was able to represent finer details of spatial information [2] or encode motion visual stimuli more faithfully than single neuron's activity [3].

It was previously reported that there are two different types of concerted firings in RGCs, with each type dependent on a particular structure of the neural circuitry – the formation of precise synchronization and correlated activity with distributed time lags were attributed to gap junctions between RGCs and common input from other neurons (amacrine/bipolar cells), respectively [4, 5].

P.-J. Liang (✉)
Department of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: pjliang@sjtu.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_27,

In the present study, the concerted pattern and its dependence on inter-cellular distance were investigated. It was observed that precisely synchronized activities and correlated activities with distributed time lags both existed among frog RGCs. Nearby cells often fired in synchrony, while remote cells were mainly occupied in correlated activities.

2 Methods

2.1 Electrophysiology Recording

Spikes from bullfrog RGCs were recorded by MEA electrodes (8 × 8, MMEP-4, CNNS UNT, USA) using a commercial multiplexed data acquisition system with a sampling rate of 40 kHz (MEA workstation, Plexon Inc. Texas, USA). A small piece (4 × 4 mm^2) of isolated retina was placed on the MEA and superfused with oxygenated (95% O_2 and 5% CO_2) standard solution, which contained (in mM): NaCl 100.0, KCl 2.5, $MgCl_2$ 1.6, $CaCl_2$ 2.0, $NaHCO_3$ 25.0, glucose 10.0.

The stimulation protocols were: (1) Pseudo-random checker-board which consisted of 16 × 16 sub-squares was displayed on the computer monitor at a frame refresh rate of 20 Hz. Each sub-square covered an area of 66 × 66 μm^2 on the retinal piece and was assigned a value either "1" (white light, 77.7 nW/cm^2) or "0" (dark) following a pseudo-random binary sequence; (2) Full-field sustained white light (38.9 nW/cm^2) was given for 30 s before the stimulation protocol was applied, the purpose of which was to adjust the sensitivity of the RGCs to similar levels. Stimuli consisting of full-field white light (77.7 nW/cm^2) with duration of 1 s and dark interval of 9 s were given repeatedly for 30 times.

2.2 Data Analysis

Generally, the cross-correlation function is defined as:

$$c_{xy}(m) = \begin{cases} \dfrac{\sum_{n=0}^{N-|m|-1} x_n y_{n+m}}{R} & m \geq 0 \\ c_{yx}(-m) & m < 0 \end{cases}, \quad R = \sqrt{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i^2}, \tag{1}$$

where x_n denotes the value of sequence x at moment n; y_{n+m} is the value of sequence y at moment $n + m$; $c_{xy}(m)$, by definition, represents the correlation between sequences x and y with a time lag of m, which reflects the effect of signal x exerts on signal y with a time delay m; R is the normalizing factor.

To map the receptive fields of multiple retinal ganglion cells, spike triggered average (STA) algorithm was applied [6].

3 Results

Receptive field properties of the RGCs were estimated by STA method by applying pseudo-random checker-board stimulation, and distances between any two RGCs' receptive field centers were calculated. Figure 1a, b demonstrate an example showing the receptive field spatial profiles of two RGCs recorded from one piece of retina. The distance between the receptive field centers was measured 334.57 μm.

Figure 2 illustrates two types of concerted pattern among frog RGCs in response to full-field white light flashes. Figure 2a shows the example of precisely synchronized activities (peak width = 1 ms) between two adjacent RGCs (distance = 61.82 μm); Figure 2b shows the example of correlated activities with distributed time lags (peak width = 40.67 ms) between two remote RGCs (distance = 365.02 μm).

Further analysis performed on 253 pairs of neurons showed that the concerted pattern varied with the distance between the two neurons' receptive field centers. Figure 3 gives the summarized results. At short distance (< 250 μm), precise synchronization was the major component. Over longer distance (> 250 μm), correlation with distributed time lags was generated more often. These results suggest that nearby neurons tended to fire in precise synchrony, while remote cells were occupied in correlation with distributed time lags more often.

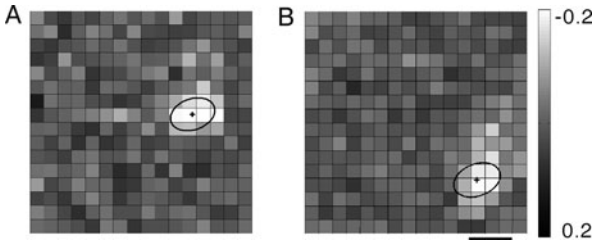


Fig. 1 The receptive field profiles of two RGCs. **a** and **b** Receptive field map of two neurons respectively. The crosses indicate the receptive field center of each neuron, the distance between which was measured 334.57 μm; 1-s.d. boundary is also indicated (scale bar, 200 μm)

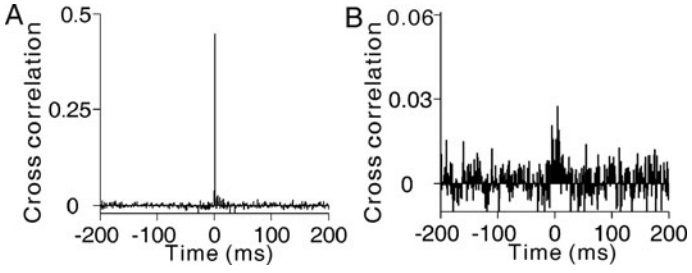


Fig. 2 Examples of two different concerted patterns among bullfrog RGCs in response to full-field *white* light flashes. **a** Sychronized activities. **b** Correlated activities

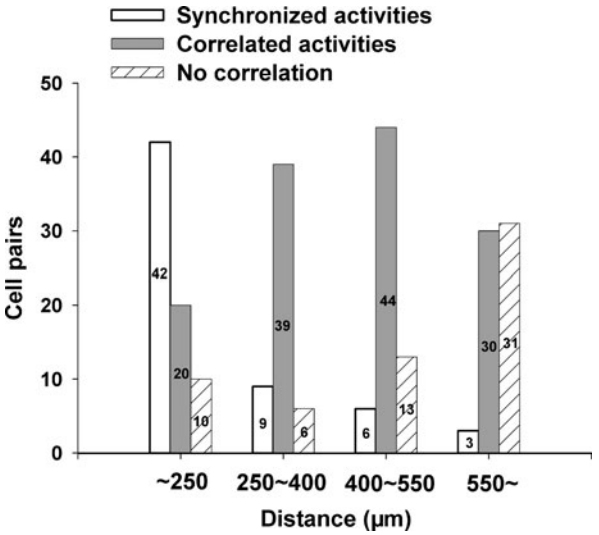


Fig. 3 Concerted pattern varied with the distance between the two neurons' receptive fields

4 Discussion

In our experiments, two types of concerted activities – synchronized activities and correlated activities with distributed time lags – were observed. Concerted activity pattern of RGCs varied with the distance between the neurons, which is related to the anatomy characters of retina. Adjacent RGCs are more likely to be connected by gap junction and generate precise synchronization more often. Correlated activity between two remote RGCs was more likely to involve the activity of amacrine cells or bipolar cells. Thus, remote cells often fired in correlation with distributed time lags.

Acknowledgments This work was supported by grants from Hi-Tech Research and Development Program of China (No.2006AA01Z125) and National Foundation of Natural Science of China (No. 30670519).

References

1. Barlow, H.B.: The Ferrier lecture, 1980. Critical limiting factors in the design of the eye and visual cortex. Proc. R. Soc. Lond. B Biol. Sci. **212** (1981) 1–34.
2. Schnitzer, M.J., Meister, M.: Multineuronal firing patterns in the signal from eye to brain. Neuron. **37** (2003) 499–511.
3. Frechette, E.S., Sher, A., Grivich, M.I., Petrusca, D., Litke, A.M., Chichilnisky, E.J.: Fidelity of the ensemble code for visual motion in primate retina. J. Neurophysiol. **94** (2005) 119–135.
4. Brivanlou, I.H., Warland, D.K., Meister, M.: Mechanisms of concerted firing among retinal ganglion cells. Neuron. **20** (1998) 527–539.

5. Meister, M., Lagnado, L., Baylor, D.A.: Concerted signaling by retinal ganglion cells. Science. **270** (1995) 1207–1210.
6. Meister, M., Pine, J., Baylor, D.A.: Multi-neuronal signals from the retina: acquisition and analysis. J. Neurosci. Methods. **51** (1994) 95–106.

Gamma-Frequency Synaptic Input Enhances Gain Modulation of the Layer V Pyramidal Neuron Model

Xiumin Li, Kenji Morita, Hugh P.C. Robinson, and Michael Small

Abstract Cortical gamma frequency (30–80 Hz) oscillations have been suggested to underlie many aspects of cognitive functions. In this paper we compare the $f - I$ curves modulated by gamma-frequency-modulated stimulus and Poisson synaptic input at distal dendrites of a layer V pyramidal neuron model. The results show that gamma-frequency distal input amplifies the sensitivity of neural response to basal input, and enhances gain modulation of the neuron.

1 Introduction

Gamma frequency oscillation, which is thought to be an emergent phenomenon generated by the synchronization of inhibitory fast spiking interneuron networks [1–4], plays a role in sensory binding, attentional selection, and consciousness. Particularly, this rhythmic inhibitory synchronization entails rhythmic gain modulation [5]. Gain modulation is defined as the change of the tuning curve of a neuron to one input as a second input is modified. Gain control has been studied by turning the membrane depolarization, noise intensity [6] and network properties [7]. The effect of gamma frequency inhibitory synchrony on gain modulation in a simple Hodgkin-Huxley neuron model was investigated in [8].

In this paper, the $f - I$ curve of a detailed layer V pyramidal neuron model is studied, where synaptic inputs are received from both basal and distal dendrites of the neuron. We examined how the relationship between the basal input and output firing rates is affected by the distal input. Results show that gain modulation occurs and gamma-frequency distal input has a stronger gain modulation than the Poisson distal input. Meanwhile the output firing becomes gamma-modulated and its spike phase distribution closely matches the experimentally-obtained spike timing density curve [1].

X. Li (✉)
Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong
e-mail: freexmin@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_28,

2 Methods

Simulations were performed using NEURON version 7.0 [9]. The layer V neocortical pyramidal neuron model was obtained from the ModelDB section of the Senselab database (http://senselab.med.yale.edu). This realistic cell model includes five voltage-dependent currents: fast $Na^+(I_{Na})$, fast $K^+(I_{Kv})$, slow, non-inactivating $K^+(I_{Km})$, and high-threshold $Ca^{2+}(I_{Ca})$, as well as one Ca^{2+}-dependent potassium current (I_{KCa}). The time step was 50 μs. Values of channel densities are given in Table 1. Each pre-synaptic input was modeled by using AMPA and NMDA conductances for excitation, GABA conductance for inhibition. The AMPA and GABA conductance was implemented using the built-in function of NEURON (Exp2Syn()). The NMDA conductance was implemented based on the asymmetric trapping block model of NMDA receptors [10]. All activated synapses were uniformly distributed at the distal and basal dendrites. Data was recorded from the last 7,000 ms of the simulations.

3 Results

Figure 1 shows the structure of the neuron model and its stimulus: (1) distal modulative input, which is composed of Poisson excitatory current and inhibitory (GABA) current; (2) basal driving input, which is composed of only Poisson excitatory current. The ratio between excitation and inhibition is fixed to be 4:1. The mean firing frequency of each pre-synaptic neuron is 10 Hz for the excitatory regular spiking (RS) pyramidal neurons and 40 Hz for the fast spiking (FS) inhibitory interneurons. The gamma GABA current was generated by the partially synchronized FS neurons, whose spiking times follow the observed in vivo phase distribution of FS cellular spikes (Fig. 3a, c: blue dotted line) [1, 2].

Figure 2a shows that both gamma and Poisson distal input can give an obvious change in the shape of the $f - I$ curve. The general firing rate to gamma input is higher than the Poisson input especially when the basal input is weak. With the increase of the number of distal synapses, the difference of corresponding firing frequency between gamma and Poisson input is enlarged due to the rapid increase of the variance of gamma input (Fig. 2b). From Fig. 3a, c we can see that the gamma rhythm of the membrane potential for gamma input can be transferred to the soma. This periodic inhibitory input produces more fluctuations at the distal dendrites than the random input (Fig. 3b, d). Thus it gives more opportunities for the neuron to generate action potentials and results in a higher firing frequency than the random input. Due to the large variance of periodic current, the large fluctuations at distal dendrites could prevent the neuron from reaching saturation. Hence, the gamma distal input generated by synchronized FS cells can give a wider range of modulation for the basal input-output relationship than the random input. Further, the spike phase distribution of this RS pyramidal cell is consistent with the experimentally observed

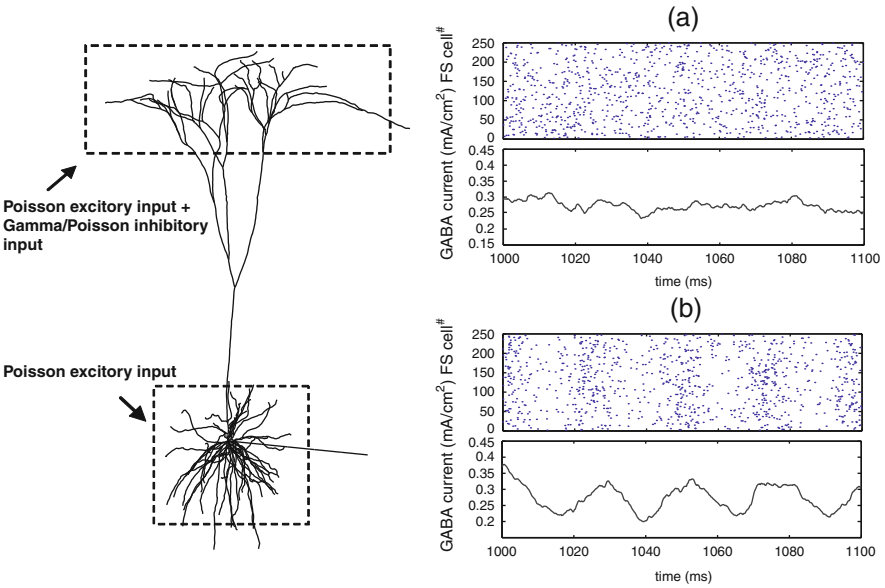


Fig. 1 *Left*: structure of the layer V pyramidal neuron model; *Right*: poisson (**a**) and gamma (**b**) GABA input current (outward) which is respectively generated by non-synchronized and partially synchronized FS inhibitory neurons (Colors will be viewed in online version)

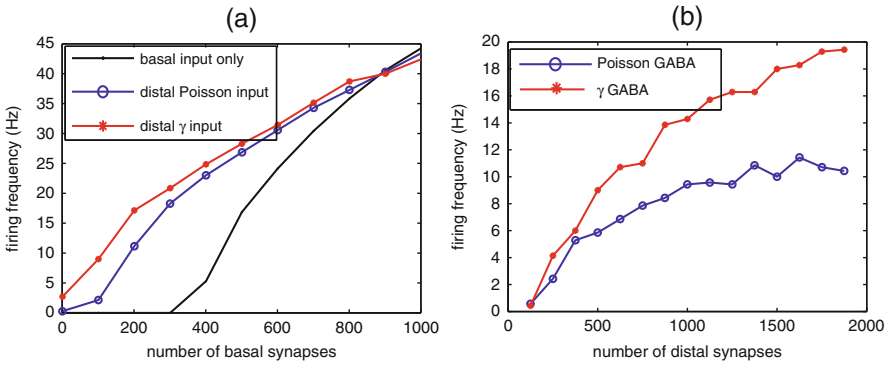


Fig. 2 **a** Gain modulation of distal stimulus for three cases: only basal input (*black*), Poisson GABA distal input (*blue*) and gamma GABA distal input (*red*); **b** Firing rate at soma with the increase of total number of distal synapses (excitation + inhibition); here the number of basal synapses is 200 (Colors will be viewed in online version)

spike density of gamma-modulated RS cells [1]. This result further confirms that RS cells are likely to receive strong gamma-modulated inhibition from FS cells and excitation with little or no gamma modulation from other RS cells [2].

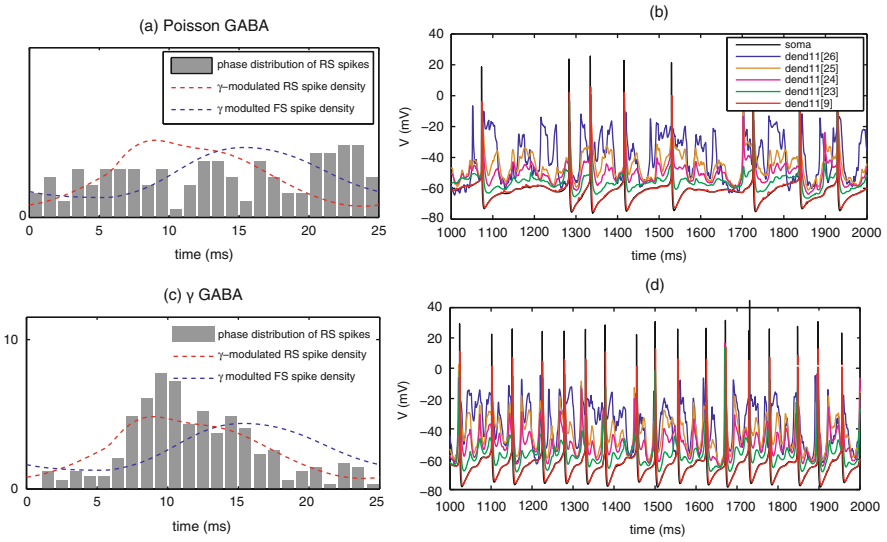


Fig. 3 Comparing the responsiveness to poisson (**a**) (**b**) and gamma (**c**) (**d**) distal input. In both cases, the number of excitatory and inhibitory synapses is 1,000 and 250 respectively. (**a**) (**c**) Histograms of the spike phases for this neuron with poisson GABA input (*left*) or gamma GABA input (*right*). *Red line*: gamma-modulated spike density for RS cells; *blue line*: gamma-modulated spike density for FS cells; (**b**) (**d**) Recordings of membrane potentials at soma and dendrites

4 Conclusion

In this paper we investigate how the gamma-frequency inhibitory stimulus at distal dendrites (modulative input) together with Poisson stimulus at basal dendrites (driving input) affect the somatic firing frequency of a layer V pyramidal neuron model. Results show that gamma-frequency distal input can give a stronger gain modulation than the Poisson distal input, and make the somatic output firing become gamma-modulated.

References

1. Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., McCormick, D.: Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. Neuron **47**(3) (2005) 423–435.
2. Morita, K., Kalra, R., Aihara, K., Robinson, H.: Recurrent synaptic input and the timing of gamma-frequency-modulated firing of pyramidal cells during neocortical "UP" states. J. Neurosci. **28**(8) (2008) 1871.
3. Cardin, J., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L., Moore, C.: Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature (2009).
4. Sohal, V., Zhang, F., Yizhar, O., Deisseroth, K.: Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature (2009).

5. Fries, P.: Neuronal gamma-band synchronization as a fundamental process in cortical computation. Ann. Rev. Neurosci. **32**(1) (2009).
6. Shu, Y., Hasenstaub, A., Badoual, M., Bal, T., McCormick, D.: Barrages of synap-tic activity control the gain and sensitivity of cortical neurons. J. Neurosci. **23**(32) (2003) 10388–10401.
7. Brozović, M., Abbott, L., Andersen, R.: Mechanism of gain modulation at single neuron and network levels. J. Comput. Neurosci. **25**(1) (2008) 158–168.
8. Tiesinga, P., Fellous, J., Salinas, E., José, J., Sejnowski, T.: Inhibitory synchrony as a mechanism for attentional gain modulation. J. Physiol. (Paris). **98**(4–6) (2004) 296–314.
9. Hines, M., Carnevale, N.: The NEURON simulation environment. Neural. Comput. **9**(6) (1997) 1179–1209.
10. Vargas-Caballero, M., Robinson, H.: Fast and slow voltage-dependent dynamics of magnesium block in the NMDA receptor: the asymmetric trapping block model. J. Neurosci. **24**(27) (2004) 6171–6180.

Zinc Modulation of Calcium-Permeable AMPA Receptors on Carp Retinal Horizontal Cells

Yan Sun, Xiao-Dong Jiang, Lei Xiao, Hai-Qing Gong, and Pei-Ji Liang

Abstract Calcium-permeable AMPA receptors are proved to be expressed on carp retinal horizontal cells. However, the synaptic contribution and modulatory characteristics of these receptors in the synapses between cones and horizontal cells are not fully understood. In the present study, we used whole-cell patch clamp technique to test the functional contribution of calcium-permeable AMPA receptors and relevant zinc modulatory effect on these receptors. It is shown that application of 100 μM pentobarbital (a selective blocker of calcium-impermeable AMPA receptors) slightly but significantly suppressed the current elicited by 3 mM glutamate. Furthermore, the co-application of 300 μM zinc apparently inhibited the remaining calcium-permeable AMPA receptors. These results suggest that calcium-permeable AMPA receptors contribute a large portion to the cell's glutamate-current and zinc can down-regulate the function of these receptors.

Keywords Glutamate current · Patch-clamp recording · Pentobarbital · Zinc

1 Introduction

AMPA receptors (AMPARs) mediate the majority of fast excitatory neurotransmission at glutamatergic synapses in the central nervous system, and can be divided into calcium-permeable and calcium-impermeable subtypes [1].

Horizontal cells (HCs) are second-order interneurons in the inner nuclear layer of vertebrate retina and respond to glutamate released from photoreceptors. Their activities regulate the visual signals laterally and determine the antagonistic

P.J. Liang (✉)
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: pjliang@sjtu.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_29,

receptive field structure of bipolar and ganglion cells. It is well accepted that ionotropic glutamate receptors expressed on carp retinal HCs are AMPA-preferring and our previous work revealed that they are partly of calcium-permeable subtype (CP-AMPARs) [1].

In vertebrate retina, zinc is co-localized with presynaptic glutamate vesicles and released from photoreceptor terminals, which suggests that zinc may play a modulatory role in the outer retina [2]. Moreover, experimental results show that the effects that zinc exerts on glutamate receptors in vertebrate HCs are diversified [3].

Here, we examined the functional contribution of CP-AMPARs expressed on carp retinal HCs and the zinc modulatory effect on these AMPARs. The application of specific calcium-impermeable AMPAR (CIP-AMPAR) inhibitor, 100 μM pentobarbital, effectively but moderately suppressed the glutamate-elicited current, suggesting that CP-AMPARs mediate the majority of the total glutamate current. Furthermore, co-application of zinc (300 μM) attenuated CP-AMPAR-mediated current. This implies that zinc down-regulates the glutamate-response through CP-AMPAR in carp retinal HCs.

2 Methods

2.1 Preparation

The experiments were performed on H1 HCs isolated from adult carp (*Carassius Auratus*, 15–20 cm body length) retinas, following the method previously described [4]. H1 cell was easily distinguished by its characteristic morphology under microscope.

2.2 Whole-Cell Recording and Drug Application

Cells were voltage-clamped at −60 mV and whole-cell recordings were achieved by 5–8 MΩ patch pipette pulled from borosilicate glass (Sutter Instrument Inc., USA) using a horizontal puller (P87, Sutter Instrument Inc., USA). The pipette was filled with intracellular solution, mounted on a motor-driven micromanipulator (MC1000e, SD Instrument Inc., USA), and was connected to a patch amplifier (Axopatch 200B, Axon Instrument Inc., USA). Data acquisition was performed using AxoScope software (Axon Instrument Inc., USA), with sample rate being 1 kHz. The recorded data were analyzed by Clampfit 9.2 software (Axon Instrument Inc., USA). All the drugs were briefly applied using the superfusion system (DAD-12, ALA Scientific Instruments, USA). The statistical data are all presented in the form of mean ± SD in the text and mean ± SE in the figure illustrations. Paired t-test was performed for statistical analysis.

3 Results

3.1 The CP-AMPARs Mediated the Majority of the Glutamatergic Response

Pentobarbital (PB), at a concentration of 100 μM, has been reported to be a selective antagonist of CIP-AMPARs [5]. So we used 100 μM PB to isolate the CP-AMPAR-mediated response (Fig. 1). During control, the application of 3 mM glutamate elicited an inward current in an H1 cell with peak value of 515.6 pA. After 10 s pre-superfusion of 100 μM PB, the peak value was reduced to 380.1 pA. The glutamate-response was almost completely recovered (peak current 494.3 pA) when PB was washed out for 15 s. The statistical results given in Fig. 1 show that the peak value of the glutamate-response was reduced to 73.2 ± 3.1% of the control level after application of 100 μM PB ($n = 5,\ p < 0.05$).

3.2 Zinc Inhibitory Effect on CP-AMPAR-Mediated Current

To investigate the zinc modulatory effect on glutamate-response mediated by CP-AMPARs, 100 μM PB was used. In the presence of 100 μM PB, application of 3 mM glutamate elicited a peak current of 427.9 pA (Fig. 2). After pre-superfusion of 300 μM zinc for 10 s, co-application of 3 mM glutamate elicited a peak current of 214.6 pA. This current was mostly recovered after zinc was washed away for more than 1 min (peak current 408.8 pA). Statistical data show that the application of 300 μM zinc reduced the CP-AMPAR-mediated peak current to 53.4 ± 3.8% of the control level ($p < 0.05,\ n = 5$).

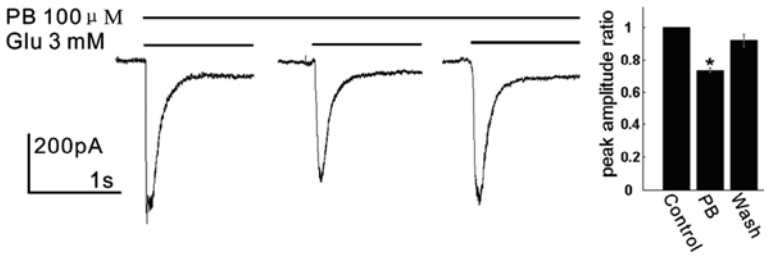


Fig. 1 Inward current elicited by 3 mM glutamate recorded from an H1 cell was slightly reduced by 100 μM PB and recovered on washout. Normalized data of PB inhibition (Mean ± SE) are plotted in the *right panel* ($n = 5$, $^*p < 0.05$, paired *t*-test)

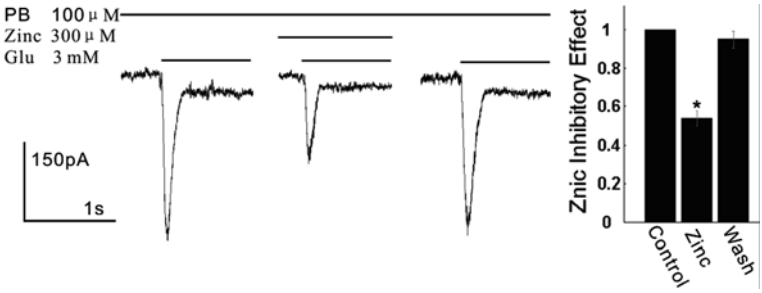


Fig. 2 An inward current evoked by 3 mM glutamate in the presence of 100 μM PB recorded from one H1 cell was suppressed by 300 μM zinc and mostly recovered on washout. Normalized zinc inhibitory effect (Mean ± SE) on currents mediated by CP-AMPARs ($^*p < 0.05$, paired *t*-test, $n = 5$)

4 Summary

The main results of this study are: (1) in isolated carp retinal H1s, CP-AMPARs mediate the majority (about 70%) of the cell's glutamate-response; (2) zinc, at a concentration of 300 μM, effectively down-regulates the CP-AMPAR-current.

The expression of CP-AMPARs in HCs is important to the regulation of synaptic plasticity between photoreceptors and HCs [6]. During dark, the activation of CP-AMPARs depolarizes the membrane potential and motivates GABA release. However, calcium influx into the cell through CP-AMPARs also triggers the intracellular calcium pathways and the increasing of intracellular calcium would inhibit the GABA transports [4]. The zinc effect of CP-AMPARs therefore mediates the local fine modulation for the reciprocal synapse between photoreceptors and HCs (Fig. 3). In addition, CP-AMPARs also play a pathological role in excitotoxicity caused by retina ischemia due to their high calcium permeability; therefore the

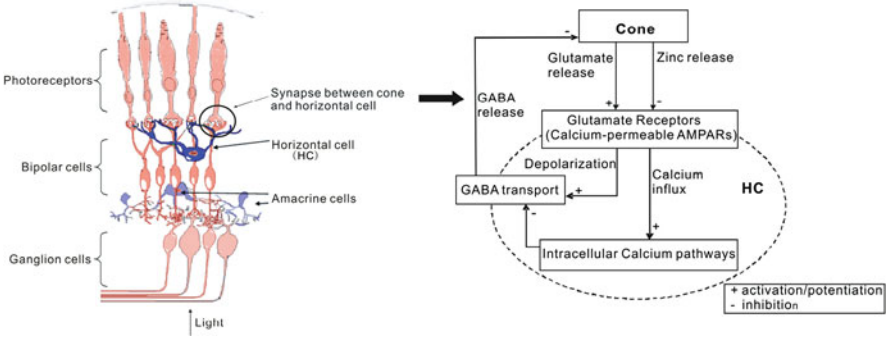


Fig. 3 The retina structure (*left panel*) and the synaptic connection between cone photoreceptors and *horizontal cells* (*right panel*). The *right panel* briefly illustrates the possible involvement of CP-AMPARs in the modulation of synaptic activities between cone and *horizontal cell*

inhibitory effect that zinc exerts on CP-AMPARs may probably protect the neurons from cell death [3].

Acknowledgments This work was supported by grants from National Foundation of Natural Science of China (No. 30870836) and Instrumental Analysis Center of SJTU.

References

1. Huang, S.Y., Liang, P.J.: Ca^{2+}-permeable and Ca^{2+}-impermeable AMPA receptors coexist on horizontal cells. Neuroreport. **16** (2005) 263–266.
2. Redenti, S., Ripps, H., Chappell, R.L.: Zinc release at the synaptic terminals of rod photoreceptors. Exp. Eye Res. **85** (2007) 580–584.
3. Zhang, D.Q., Ribelayga, C., Mangel, S.C., McMahon, D.G.: Suppression by zinc of AMPA receptor-mediated synaptic transmission in the retina. J. Neurophysiol. **88** (2002) 1245–1251.
4. Jiang, X.D., Wang, X.L., Sun, Y., Gong, H.Q., Liang, P.J.: NMDA modulation of GABA transporter current in carp retinal horizontal cells. Brain Res. **1240** (2008) 105–110.
5. Taverna, F.A., Cameron, B.R., Hampson, D.L., Wang, L.Y., MacDonald, J.F.: Sensitivity of AMPA receptors to pentobarbital. Eur. J. Pharmacol. **267** (1994) R3–R5.
6. Huang, S.Y., Hu, J.F., Gong, H.Q., Liang, P.J.: Postsynaptic calcium pathway contributes to synaptic plasticity between retinal cones and luminosity-type horizontal cells. Sheng Li Xue Bao. **58** (2006) 407–414.

Spectral Characteristics of Firing Pattern in Retinal Ganglion Cells

Hao Li and Pei-Ji Liang

Abstract This paper studies the firing patterns of retinal ganglion cells (RGCs) with frequency-domain approach based on partial directed coherence (PDC) which reflects a frequency-domain representation of Granger causality. The results suggest that the PDC of the RGCs' spike trains in response to natural movie has significant frequency dependence; while it doesn't change much across neighboring frequency band under the pseudo-random checker-board stimulus. Possible mechanisms underlying these distinct features are discussed.

Keywords Checker board · Natural movie · Granger causality · Partial directed coherence

1 Introduction

Neural codes for visual information have probably evolved to maximize their efficiency of metabolic activity or information transmission [1]. Thus it is reasonable to assume that the early stage of visual information processing is to encode incoming stimuli to improve efficiency [1, 2].

Retina is the first stage of visual neural information processing. We are interested in the population coding scheme that RGCs adopt to ensure their coding efficiency. By using multi-electrode array system, activities of RGCs in response to different stimuli were recorded and analyzed.

H. Li (✉)
Biomedical Engineering Department, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: livio.sjtu@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_30,

2 Methods

2.1 Electrophysiology Recording

Experiments followed the procedures described by Liu et al. [3]. Briefly, retina was isolated from bullfrog under dim red illumination and perfused using oxygenated Ringer's solution. The retina was dissected into 3×3 mm^2 squares. One small piece of retina was placed with the ganglion cell side down in a perfusion chamber the bottom of which contained a microelectrode array (MEA). The MEA consisted of 60 electrodes arranged in an 8×8 matrix (leaving the 4 corners void). The tissue and perfusate were kept at room temperature.

2.2 Stimulus

Visual stimulation included (1) natural movies and (2) pseudo-random checkerboard. The stimuli were displayed at a rate of 20 frames per second. A 5-s segment of stimuli was repeated 20 times with interleaving 5-s full-field gray screen (grayscale = 128).

2.3 Power Spectral Density (PSD)

Peri-stimulus time histogram was accumulated from 20 trials of spike trains. The power spectral density (PSD) function of the PSTH was calculated using the standard fast Fourier transform algorithm.

2.4 Partial Directed Coherence (PDC)

The concept of Granger-causality, originated from econometrics [4], is a fundamental tool for the description of directed dynamic relationships among the components of a multivariate process and has been applied recently to problems in neuroscience. Based on the common sense conception that causes precede their effects in time, this probabilistic concept of causality exploits the temporal structure of signals and defines causal relationships in terms of predictability. In a linear framework, the notion of Granger-causality is closely related to multivariate autoregressions. In this paper, we considered the PSTHs from n RGCs as an n-dimensional time series, which was used to construct a multi-variant auto-regressive (MVAR) model.

Let $X = X(t), t \in \mathbb{Z}$ with $X(t) = (x_1(t), \ldots, x_n(t))'$ be the PSTHs of n RGCs. Then an MVAR model of order p for X is given by

$$X(t) = \sum_{r=1}^{p} a(r)X(t-r) + \varepsilon(t) \tag{1}$$

where $a(r)$ are the $n \times n$ coefficient matrices of model and $\varepsilon(t)$ is a multivariate Gaussian white noise process.

In order to provide a frequency domain description of Granger-causality, Baccalá and Sameshima [5] introduced the concept of partial directed coherence (PDC). Let

$$A(f) = I - \sum_{r=1}^{p} a(r)e^{-i2\pi fr} \tag{2}$$

denote the difference between the n-dimensional identity matrix I and the Fourier transform of the coefficient series. Then the partial directed coherence $\left|\pi_{i\leftarrow j}(f)\right|$ is defined as

$$\left|\pi_{i\leftarrow j}(f)\right| = \frac{\left|A_{ij}(f)\right|}{\sqrt{\sum_k \left|A_{kj}(f)\right|^2}} \tag{3}$$

where $\left|\pi_{i\leftarrow j}(f)\right|$ denotes the causal influence from the past of time series $x_j(t)$ to the present of $x_i(t)$.

3 Results

3.1 Power Spectral Density of Responses of RGCs

In the present study, we first calculated the PSD of PSTH in response to checker-board and natural movie respectively. Figure 1 shows the power spectra of 36 RGCs' (the positions of which are marked with circles in Fig. 2) firing activities.

It is shown that the power of low frequency components in movie-responses was stronger than that of checker-board-responses with a slower decay.

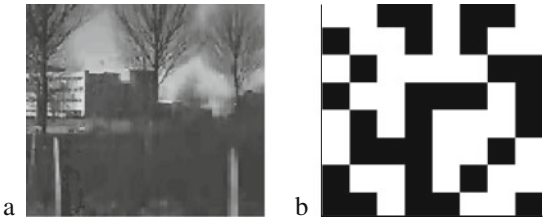


Fig. 1 Visual stimuli used: **a** Example movie frame (http://hlab.phys.rug.nl/vidlib/index.html), mean *grayscale* ranging from 99 to 117 in *grayscale*. **b** Example checker-board frame (8×8 grid, *grayscale*: *black* = 0, *white* = 255)

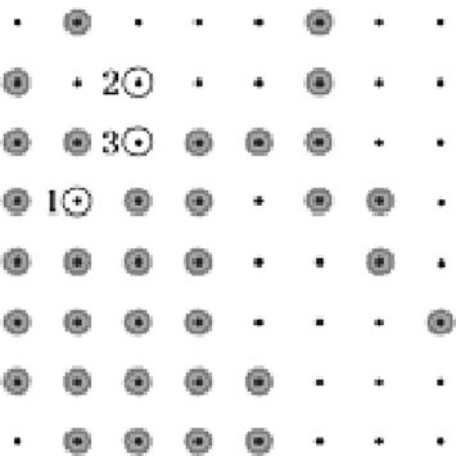


Fig. 2 Locations of the 36 RGCs recorded from a retina, and the 3 selected ones

3.2 Partial Directed Coherence Analysis

RGCs together with their presynaptic neurons form a complicated retinal circuit. Outputs of different RGCs exhibit more or less interactive influences probably caused by the presynaptic input. We therefore analyze the spike trains fired by RGCs and expect to acquire interactive relationship of spike trains.

Figure 3 illustrates the PDC calculated for 3 RGCs (open circles in Fig. 2). It is shown that the PDC of the movie-responses oscillated significantly and presented sharp peaks in different frequencies, while the PDC of the checker-board-response was less changeable. This might indicate that the relationships between the RGCs' activities during the natural movie are more complicated than that during the checker-board, and suggests that influence between outputs of RGCs is more intensive in response to the natural movie.

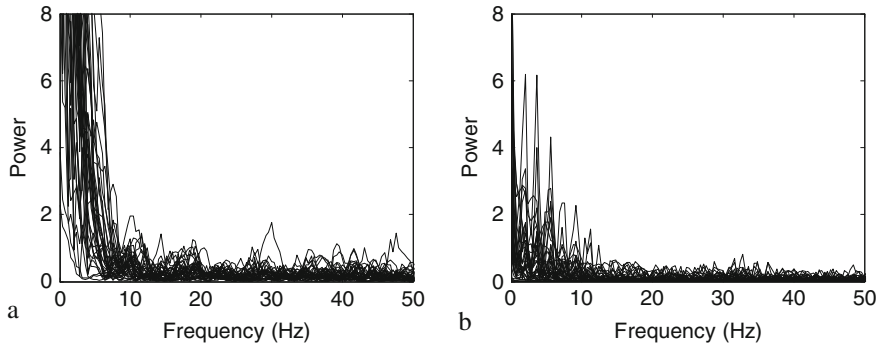


Fig. 3 Summary of power spectral density of the 36 RGCs (*circles* in Fig. 2) in response to natural movie (**a**) and checker-board (**b**)

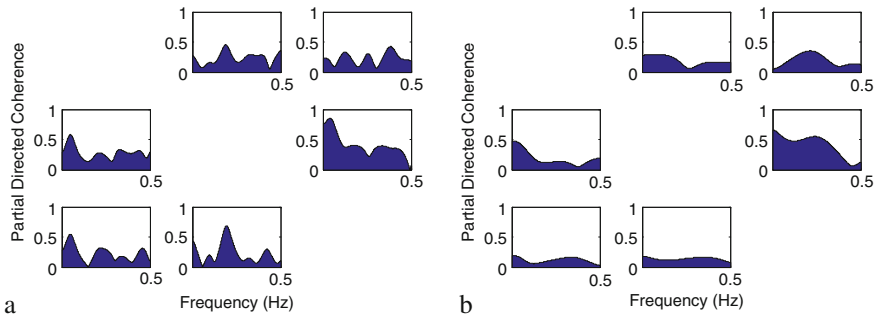


Fig. 4 PDC analysis: **a** PDC of 3 RGCs (marked by *open circles* in Fig. 2) in response to natural movie. **b** PDC of the same RGCs in response to checker-board. ($i, j = 1, 2, 3$)

4 Discussion

The PSD analysis results described here suggest that firing patterns of RGCs in response to natural movie and checker-board have distinct characteristics, and the distinction may attributes to the different presynaptic inputs.

The causal influence between neurons, indicated by the PDC analysis, are probably not caused by direct interactions between the neurons being investigated (the distance between the adjacent microelectrodes is 100 μm), but modulated by the presynaptic inputs. Retinal neurons, especially bipolar cells and amacrine cells, form a neural circuit to provide presynaptic signals to RGCs. Presynaptic signals conducted through different pathways will combine with different synaptic delays. These delays might cause spike trains from different RGCs show the feature of causality.

Acknowledgements This work was supported by grants from the Hi-Tech Research and Development Program of China (No. 2006AA01Z125) and National Foundation of Natural Science of China (No. 30670519).

References

1. Barlow, H.B.: Unsupervised learning. Neural Comput. **1** (1989) 295–331.
2. Dan, Y., Atick, J.J., Reid, R.C.: Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J. Neurosci. **16**(10) (1996) 3351–3362.
3. Liu, X., Zhou, Y, Gong, H.Q., Liang, P.J.: Contribution of the GABAergic pathway(s) to the correlated activities of chicken retinal ganglion cells. Brain Res. **1177** (2007) 37–46.
4. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica. **37**(3) (1969) 424–438.
5. Baccalá, L.A., Sameshima, K.: Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. **84**(6) (2001) 463–474.

The Firing Pattern Properties in Retinal Ganglion Cells

Han-Yan Gong, Ying-Ying Zhang, Pei-Ji Liang, and Pu-Ming Zhang

Abstract The correlation between spike trains of chicken retinal ganglion cells recorded by the multi-electrode recording system are calculated by cross-correlation function and Lempel-Ziv distance. Cross-correlation evaluates the degree of neural synchronization while Lempel-Ziv distance measures the similarity of firing patterns. The results indicate that compared with synchronization, firing patterns show stronger coordination in neuron groups and reflect the stimuli property more accurately. Thus the measurement of spike train patterns by Lempel-Ziv distance can give us more comprehensive cognition and additional insights in neural coding analysis.

Keywords Firing pattern · Synchronization · Cross-correlation · Lempel-Ziv distance

1 Introduction

It is already known that information is reflected with various rules in neuron spike trains called neural coding, whose characters can be measured by different computational methods [1]. Most of algorithms addressing the correlation between spike trains to identify how neurons code information focus on comparing the precise time of coincident neural firing. Cross-correlation we used here is also one of the widely-used methods in measuring firing synchrony. However, sometimes the information carried by spike trains can not be totally reflected by concerted timing property. Lempel-Ziv distance (LZ distance) is an algorithm based on Lempel-Ziv complexity which is used to estimate the amount of information increasing along the trains [2]. LZ distance considers spike trains with similar but possibly not synchronized firing

P.-M. Zhang (✉)
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: pmzhang@sjtu.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_31,

patterns as close [3]. Generally speaking, if two spike trains are highly synchronized, their patterns will be similar as well. On the contrary, if two spike trains have similar patterns, they are not necessarily synchronized in firing time. LZ distance has been proved effective in measuring neural firing reliability and neuron classification [3, 4]. Here we use LZ distance to examine whether there is any information that can't be reflected by synchronization in neural firing activities.

2 Meterial and Methods

2.1 Experimental Procedure

Spikes from ganglion cells were recorded by a multi-electrode array using a commercial multiplexed data acquisition system (MEA, Multi Channel Systems MCS GmbH, Germany) with a sampling rate of 20 kHz. The stimulation protocols were: (1) Three pieces of digitized segments of grayscale video recording (1920 frames, 128 × 128 pixels, refresh rate being 10 Hz) of natural outdoor scenes. (2) Checkerboard (8 × 8). The images were projected onto the retina piece via an optical lens system and covered the whole area of the multi-electrode array. Each visual stimuli lasted 50 s and signals recorded by electrodes were transformed to spike trains of each individual neuron after sorting using principal component analysis and K-means clustering.

2.2 Cross-Correlation Function and LZ Distance

The normalized cross-correlation function for calculating the degree of firing correlation is defined as follows:

$$c_{xy}(m) = \begin{cases} \dfrac{\sum_{n=0}^{N-|m|-1} x_n y_{n+m}}{R} & m \geq 0 \\ c_{yx}(-m) & m < 0. \end{cases} \qquad R = \sqrt{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i^2}. \tag{1}$$

The result stands for the spike firing consistency of sequence x and y at the time delay m. The peak value represents the degree of correlation between two spike trains at the corresponding time which we choose as the evaluation value compared with LZ distance.

LZ distance is calculated as follows: The spike trains are translated into bitstrings X_n of length n using a bin size of 1 ms, such that the symbols '0' and '1' denote the absence or presence of a spike respectively. These bitstrings are coded using a parser that partitions the string into nonoverlapping substrings. As an illustration, the string 0011001010100111 is parsed as 0 |01 |1 |00 |10 |101 |001|11 . Then the Lempel-Ziv complexity $K(X_n)$ is calculated by:

$$K(X_n) = \frac{c(X_n)\log c(X_n)}{n} . \tag{2}$$

where $c(X_n)$ is the number of phrases that results from the partition of X_n. The amount of information Y_n provides about X_n is given as $K(X_n) - K(X_n|Y_n)$, where $c(X_n|Y_n)$ is the size of difference of the phrases parsed from two spike trains. The LZ distance is:

$$d(X_n, Y_n) = 1 - \min\left\{\frac{K(X_n) - K(X_n|Y_n)}{K(X_n)}, \frac{K(Y_n) - K(Y_n|X_n)}{K(Y_n)}\right\} . \tag{3}$$

A large number of similar patterns appearing in both spike trains leads to a large overlap of the sets of phrases. Thus, distances between spike trains with similar patterns are small, whereas distances between spike trains with different patterns are large [3].

3 Results

Cross-correlation and LZ distance calculation are performed on 5 pieces of retina, recording 24, 20, 50, 55, 44 ganglion cells respectively, and reach the similar conclusion. In Fig. 1, the result of one piece of retina is shown for example. Under

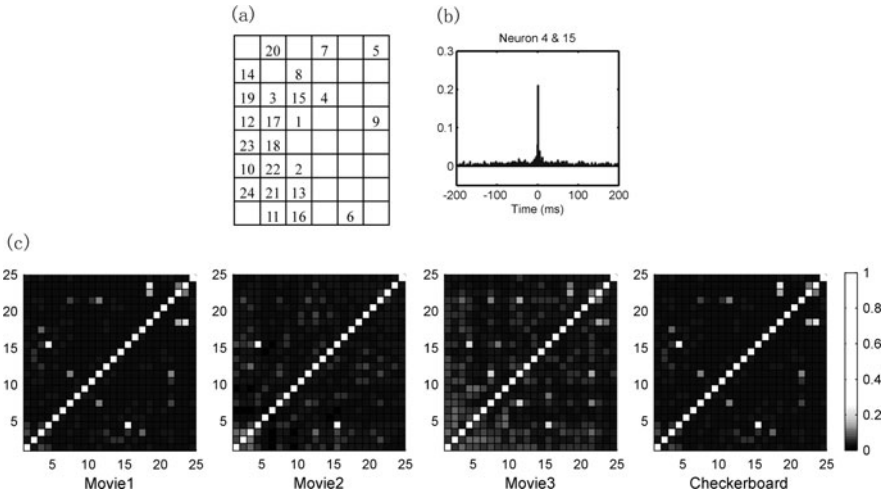


Fig. 1 The synchronized correlation between neurons measured by cross-correlation function. **a** Neurons' location where each element stands for one neuron's number (in this retina 24 neurons are recorded). **b** Cross-correlation result between two highly synchronized neurons (No.15 and No.4). **c** Synchronized correlation of neural responses under 4 stimuli. Each block stands for the cross-correlation peak value between neurons with number in x and y axis. The brighter the *grayscale* is, the more synchronized two neurons are

the stimuli of natural movies and checkerboard, overall synchronized correlation between neurons remains similar and some pairs are always highly synchronized.

Similarly, LZ distances are calculated and different collaborative features of neurons are shown in Fig. 2. Compared with the cross-correlation results, neurons show correlation in a wider range in firing patterns. Some neuron pairs which are weak in synchrony appear to be strong in firing pattern similarity, which reflect the different aspects the two algorithms focus.

The curves of cross-correlation results and LZ distance between all the neural pairs are shown in Fig. 3. Natural movies contains stronger correlation than random checkerboard whose distribution is almost the same as white noise, thus theoretically, neural response should cooperate more actively under the stimuli of movies. Both the results of cross-correlation and LZ distance meet the assumption, but the LZ distance shows more obvious difference under 4 stimuli.

In conclusion, the correlation of firing patterns between spike trains can better reflect the stimuli information coded in neural firing activities in the overall performance. Here different results are shown in measuring synchronization and firing pattern similarity, implying that firing patterns express something beyond what spike timing correlation can reflect, which will give us deeper insight in neural information coding.

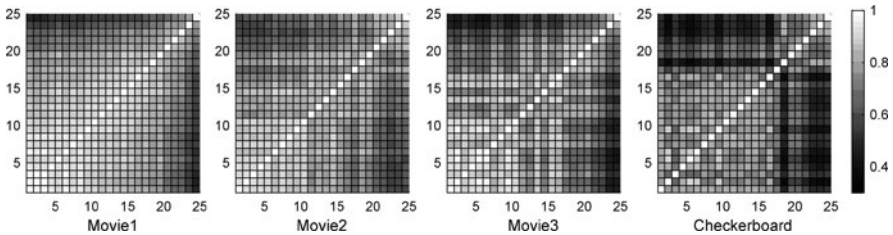


Fig. 2 The firing pattern correlation between neurons measured by LZ distance. Each block stands for the value of $1 - d(X_n, Y_n)$ between neurons with number in x and y axis. The brighter the *grayscale* is, the more similar two neurons' spike trains are

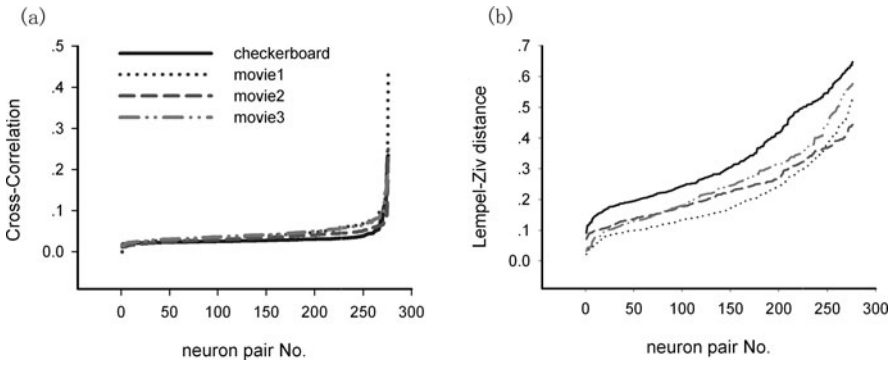


Fig. 3 The overall synchronization and firing pattern correlation of all neuron pairs

Acknowledgments This work was supported by grants from the National Natural Science Foundation of China (No. 60775034) and the Hi-Tech Research and Development Program of China (2006AA01Z125).

References

1. Rieke, F., Warland, D., de Ruyter van Steveninck, R., Bialek, W.: Spikes: exploring the neural code. Cambridge, MA: The MIT Press (1999).
2. Amigo, J.M., Szczepański, J., Wajnryb, E., Sanchez-Vives, M.V.: Estimating the entropy rate of spike trains via Lempel-Ziv complexity. Neural Comput. **16** (2004) 717–736.
3. Christen, M., Kohn, A., Ott, T., Stoop, R.: Measuring Spike Pattern Reliability with the Lempel-Ziv-Distance. J. Neurosci. Methods. **156** (2006) 342–350.
4. Christen, M., Nicol, A., Kendrick, K., Ott, T., Stoop, R.: Odour encoding in olfactory neuronal networks beyond synchronization. Neuroreport. **17** (2006) 1499–1502.

The Spatiotemporal Structure of Receptive Field of Ganglion Cells in Bullfrog Retina

Yao Li, Hai-Qing Gong, Pei-Ji Liang, and Pu-Ming Zhang

Abstract The traditional spike-triggered average method is always used to calculate the receptive fields. According to the improved spike-triggered average method, we calculated the receptive fields of ganglion cells in the bullfrog retina. The results showed that the receptive fields have dynamic spatiotemporal structure, both the response delay in the time domain and the receptive field size in the space domain possess the time-variant properties.

Keywords Receptive field · Spike-triggered average · Spatiotemporal structure · Retinal ganglion cell

1 Introduction

Retina plays an important role in processing the visual information within the early visual pathway [1]. To understand the coding mechanism that takes place in the retina, it is necessary to have a detailed understanding of the receptive field (RF) structure of the retinal ganglion cells (RGCs) [2]. The RF is classically defined as the area of visual space which can influence the discharge of a neuron. Previous studies have shown that the RF is inherently a function of both space and time [3]. Thus, to adequately describe how a neuron processes the visual image, one must characterize its RF in the space-time domain. Based on the white noise analysis [4], we can obtain the RF of the RGCs. Recent study has made improvements on the traditional spike-triggered average (STA) method [5]. Thus the spatiotemporal structure of RFs with higher time resolution can be calculated, which help define the time-variant size of the RF as well as the time-variant response delay. Physiologically, the understanding of such spatiotemporal structure of RF will be beneficial to the

P.-M. Zhang (✉)
School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: pmzhang@sjtu.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_32,

explanations of the time-varying response properties of the RGCs and facilitate the analysis of the coding mechanism.

2 Material and Methods

2.1 Experimental Procedure

The bullfrog RGCs activities were recorded simultaneously by a multi-electrode array with 60 electrodes arranged in an 8 × 8 matrix and the commercial multiplexed data acquisition system (MEA, Multi Channel Systems MCS GmbH, Germany) with a sampling rate of 20 kHz. The RGCs were stimulated with two kinds of stimuli: (1) a 24 × 24 checkerboard updated every 30 ms, lasting for 200 s; (2) a 16 × 16 checkerboard updated every 30 ms, lasting for 60 s. The spikes of the RGCs were detected by the threshold in the software MC Rack (MCS GmbH, Germany) and classified by the principal component analysis in the software Offline Sorter (Plexon Inc. Texas, USA).

2.2 Spike-Triggered Average (STA)

The spike-triggered average method is a statistical method based on the white noise analysis [4]. The checkerboard images in certain time interval before the occurrence of each spike are summed and averaged. Due to the statistical property, the stimulus patterns that elicit the spikes are strengthened and therefore reflect the RF of the neuron. The RF could be mapped according to this method, but the resolution of the RF time-variant curve is limited by the frequency of the stimulus. Recent study has made some improvements to obtain a higher time resolution [5]. Here we take into account all the checkerboard images within the 500 ms interval before each spike. The bin, which determines the resolution of the time-variance curve, is set as 10 ms. The images located in the same bin are summed and averaged. Thus there are 50 sampling points in the curve. The statistical result of each bin could efficiently describe the spatial RF structure in the corresponding time interval. By identifying the spatial coordinates of the minimum value (OFF-center ganglion cells) or maximum value (ON-center ganglion cells) in those RF maps and combining the values from all the maps in the same position together into one curve, the temporal RF structure could be calculated.

3 Results

The experiments were implemented in the bullfrog retinas and 68 ganglion cells activities in two pieces of retina were recorded. According to the improved STA method, the spatiotemporal RF structure of the RGCs with higher resolution were calculated and illustrated in Fig. 1. There are three ganglion cells with different spatiotemporal structure of RFs. The results suggest that the STA intensity reached the

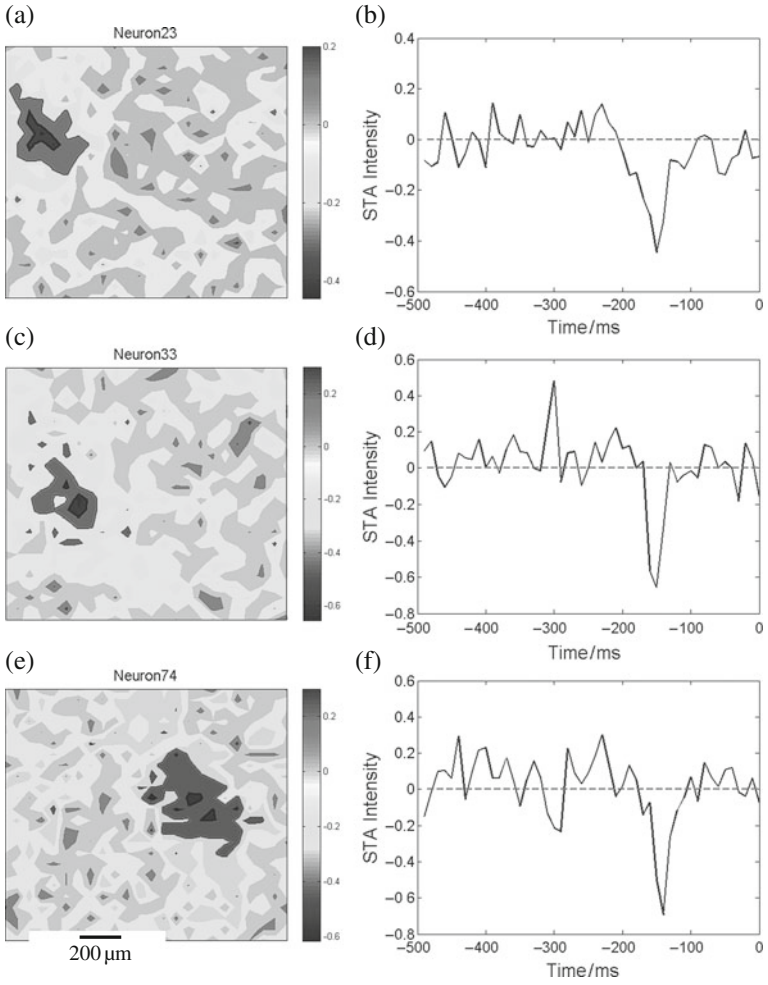


Fig. 1 The spatiotemporal RFs of different RGCs. **a** and **b** are the RF structure of the same RGC under the 24 × 24 checkerboard stimulus in the space and time domain, respectively. The spatial RF structure is corresponding to the minimum value moment of the curve in (**b**). The intensity of the area in (**a**) within the *thick lines* is below half of the minimum value in (**b**). (**c**) and (**d**), (**e**) and (**f**) are from another two ganglion cells

peak value at about 150 ms in both (b) and (d) while 140 ms in (f) before the occurrence of the spikes. The corresponding ganglion cells are all OFF-center neurons.

The RGC response delay, i.e., the time corresponding to the minimum value (OFF-center ganglion cells) or maximum value (ON-center ganglion cells) in the RF time-variant curve, as well as the size of the RF are two crucial parameters in the spatiotemporal structure. We calculated the response delay within a 60 s moving window with 10 s step resolution throughout the 200 s duration of stimulus. The results illustrated that there were two types RF of RGCs, whose response delay increased or decreased with the time respectively (Fig. 2). The numbers corresponding to the two types of RGCs were 8 and 5. Furthermore, the comparation of RF

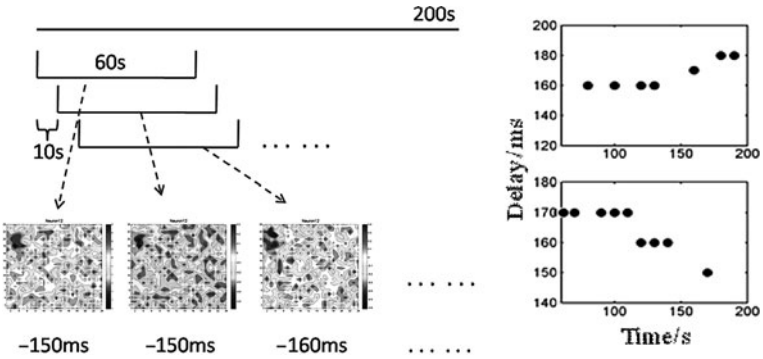


Fig. 2 The schematic of the calculation method and the time-variant RGC response delay. **a** shows the calculation method of delay curve. **b** shows the different RGCs with increasing and decreasing delay under the 24 × 24 checkerboard stimulus, respectively

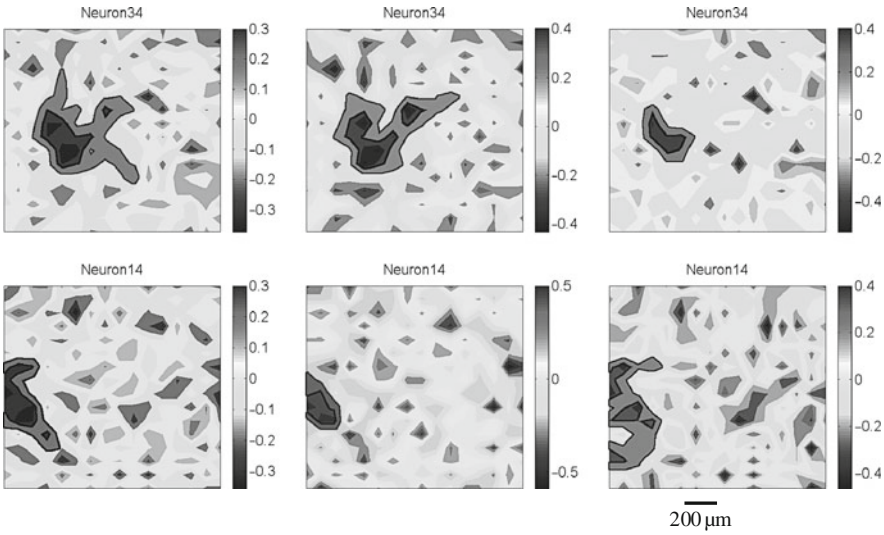


Fig. 3 The time-variant RF size. **a–c** illustrate the spatial RF calculated from total 60 s, preceding 30 s and following 30 s stimulus, respectively. The first *row* shows that the RF size of the RGC was shrinking with the time under the 16 × 16 checkerboard stimulus while the second one was reverse

size between the preceding 30 s and following 30 s during the 60 s checkerboard stimulus showed that the RF size was also time-variant, both expanding and shrinking phenomena were existed (Fig. 3). The numbers of the RGCs that possess such properties were 6 and 4, respectively.

In conclusion, the spatiotemporal structure of the RF of retinal ganglion cells with higher time resolution was obtained by using the improved STA method, which reveals the dynamic RF of the RGCs. Especially, both the response delay in the time

domain and the RF size in the space domain possess the time-variant characteristics. The future study will concentrate on how the RF dynamics makes an impact on the RGC response property as well as the coding mechanism.

Acknowledgments This work was supported by grants from the National Natural Science Foundation of China (No. 60775034) and the National Basic Research Programme of China (No. 2005CB724301).

References

1. Field, G.D., Chichilnisky, E.J.: Information processing in the primate retina: circuitry and coding. Annu. Rev. Neurosci. **30** (2007) 1–30.
2. Huberman, A.D., Feller, M.B., Chapman, B.: Mechanisms underlying development of visual maps and receptive fields. Annu. Rev. Neurosci. **31**(2008) 479–509.
3. DeAngelis, G.C., Ohzawa, I., Freeman, R.D.: Receptive-field dynamics in the central visual pathways. Trends Neurosci. **18** (1995) 451–458.
4. Meister, M., Pine, J., Baylor, D.A.: Multi-neuronal signals from the retina: acquisition and analysis. J. Neurosci. Methods. **51** (1994) 95–106.
5. Ishikane, H., Kawana, A., Tachibana, M.: Short- and long-range synchronous activities in dimming detectors of the frog retina. Vis. Neurosci. **16** (1999) 1001–1014.

Part IV
Neural Population Dynamics

Dynamics of Hierarchical Neural Networks

Claus C. Hilgetag, Mark Müller-Linow, and Marc-Thorsten Hütt

Abstract Neural networks are organized hierarchically across many scales of dimension, from cellular neuronal circuits via mesoscopic networks at the level of columns, layers and areas to large-scale brain systems. However, the structural organization and dynamic capabilities of hierarchical networks are still poorly characterized. We investigated the contribution of different features of network topology to the dynamic behavior of hierarchically organized neural networks. Prototypical representatives of different types of hierarchical networks as well as two biological neural networks were explored with a three-state model of node activation for systematically varying levels of random background network stimulation. The results demonstrated that two principal topological aspects of hierarchical networks, node centrality and network modularity, correlate with the network activity patterns at different levels of spontaneous network activation. The approach further showed that the dynamic behavior of the cortical systems network in the cat is dominated by the network's modular organization, while the activation behavior of the cellular neuronal network of *Caenorhabditis elegans* is strongly influenced by hub nodes. Generally, our results demonstrate the interaction of multiple topological features and dynamic states in the function of complex neural networks.

Keywords Complex networks · Neural topology · Modularity · Betweenness centrality · Cat · *Caenorhabditis elegans*

1 Introduction

Neural networks are organized hierarchically across many scales of dimension, from cellular neuronal circuits via mesoscopic networks at the level of columns, layers and areas to large-scale brain systems. However, the structural organization and

C.C. Hilgetag (✉)
School of Engineering and Science, Jacobs University Bremen, 28759 Bremen, Germany;
Department of Health Sciences, Boston University, Boston, MA 02215, USA
e-mail: c.hilgetag@jacobs-university.de

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_33,

dynamic capabilities of hierarchical networks are so far poorly characterized. For instance, hierarchical organization in networks can be described by several different characteristics, such as (i) the existence of network hubs with a large number of connections, (ii) repeated encapsulation of sets of nodes in increasingly larger sets, as well as (iii) the self-similarity of the network across dimensional scales. Moreover, the combination of modular and hub features can produce various types of network topologies. Classical Erdös-Rényi (ER) random graphs do not contain hubs or modules and may thus serve as a general null model. Scale-free Barabási-Albert (BA) graphs [1], on the other hand, contain only hubs and no modules. Within such graphs, projections from the hubs can reach many network regions, and the hub nodes thus have a more privileged role than nodes with fewer connections and a more local reach. Further, networks that do not contain hubs, but are modular, may arise from linking many distributed, dense clusters with a small number of inter-cluster connections. Such clusters can exist at different levels (representing clusters of sub-clusters of sub-sub-clusters [2]), resulting in a hierarchical network organization, which has recently been termed "fractal" [3]. Finally, networks may be modular and also contain hubs, which are either contained within the modules serving as local hubs, or may form global hubs that integrated network modules at different scales of organization. The two latter networks combine features of scale-free and modular networks. Currently, it is not clear how the combination of these different topological aspects may shape the dynamic behavior of hierarchical neural networks. In order to investigate this question, we modeled the propagation of spontaneous activity in different prototypes of hierarchical neural networks, as well as for two examples of biological neural networks, the cellular neuronal network of *C. elegans* and the global network of cortical systems connections in the cat.

2 Methods

In order to study dynamic patterns on different types of hierarchical excitable networks, we modeled the propagation of excitation across these networks using a discrete excitable "forest-fire" model [4], which reflects the principal neural cycle of excitation and deactivation followed by a refractory phase. This simple model (compared to, e.g., integrate-and-fire models) was chosen for two main reasons. First, it is currently unclear how much detail is required to realistically describe the interaction of excitable elements in complex networks, so it is best to start with a simple model. Second, the dynamic model should be general enough to be applicable to networks of individual neurons as well as neural populations. In the simulation, we considered the individual time series of all nodes, and for each pair of nodes computed the number of simultaneous firing events. When applied to the whole network, the resulting co-activation matrix represented the distribution pattern of excitations, which could be compared with a corresponding distribution pattern of

different topological properties, using the "dynamic modularity", Q_{dyn}. This measure expressed the agreement between the dynamic grouping by node co-activation and the topological grouping of nodes by distance to a central node (CN) or topological modularity (TM). For different hierarchical prototypes as well as two biological networks, the cortical systems connectivity in the cat and cellular neuronal connectivity of *C. elegans*, we then investigated which topological feature best explained the observed patterns of simultaneous firing events.

3 Results

We found that in different parameter regimes (characterized by the rate of spontaneous node excitations) different topological properties determined the observed synchronization patterns. Moreover, we showed that small systematic changes in the graph architecture, designed to enhance or decrease the selected topological property, were reflected in the dynamics. In a second step, we extended our study to hierarchically structured artificial graphs and then to biological networks, in order to demonstrate that the distribution patterns of excitations change dramatically when both properties are represented to different degrees in the respective graphs (Fig. 1). We also found that the dynamic behavior of the cerebral cortical systems network in the cat is dominated by the network's modular organization, while the activation behavior of the cellular neuronal network of *C. elegans* is strongly influenced by hub nodes (Fig. 2).

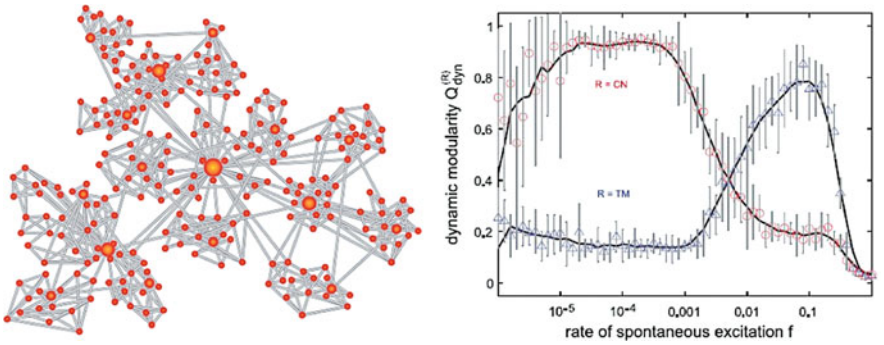


Fig. 1 *Left*: Hierarchical (scale-free modular) network combining different topological features. *Right*: Correlation of distance to central node (CN) or topological modularity (TM) with the dynamic grouping of nodes at different levels of spontaneous network activation. The dynamic modularity Q_{dyn} for both the TM reference (△ marks) and the CN reference (○ marks) is depicted as a function of the rate of spontaneous excitations *f*. The hierarchical scale-free graph displays properties of both, the modular and the BA graph. Thus, the two levels of dynamic integration are visible within the same network for the respective values of *f*. The transition between these two levels corresponds to the transition from spike to burst dynamics

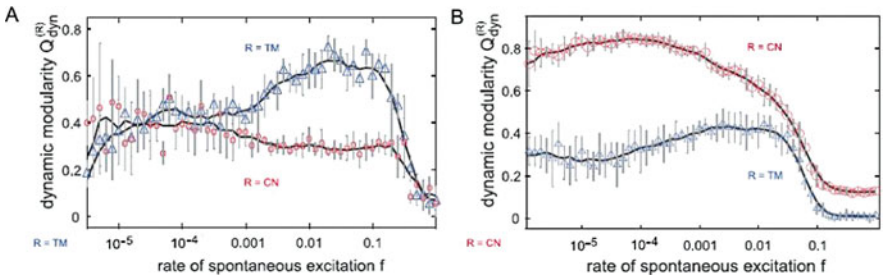


Fig. 2 Organization of dynamic behavior in two biological neuronal networks, the cortical systems network of the cat **(a)** and the cellular neuronal network of *C. elegans* **(b)**. **a** The dominance of the modular topology of the cat cortical network is reflected in a distinct increase of Q_{dyn} for the TM-dependent correlation in the high-*f* regime (△ marks), while distance from central hub nodes appears to play only a marginal role (see slight superelevation in the low-*f* regime, ○ marks). **b** By contrast, the cellular network of *C. elegans* displays a strong dependency on two adjoining central nodes which dominate the dynamics in a wide range of *f*. Particularly pronounced is the drastic increase of the CN-dependent correlations for Q_{dyn} in the low-*f* region (○ marks). Even so, there exists a noticeable but comparatively subordinate influence of the module-based excitation patterns (△ marks). Adapted from [5]

4 Discussion

The current work presents aspects of a pattern-based computational approach for linking network topology and dynamics. This approach proved useful in probing the functional organization of complex neural networks. The comparison of topological features and simulated network dynamics demonstrated that features such as central hub nodes and network modularity can strongly and systematically shape a network's dynamic behavior. Moreover, in hierarchical modular networks, where multiple of these features were present, the network dynamics exhibited a functional switch for different levels of spontaneous network activation between the dynamic organization through a central node or through modular features.

The method also revealed the dynamic impact of different topological characteristics in biological neural networks. In particular, the dynamics in the cellular neuronal network of *C. elegans* appeared organized by the topological distance to a central hub node, whereas the dynamic behavior of the cat cerebral cortical network appeared more strongly influenced by network modularity. Both topological features, however, contribute to the organization of the networks synchronization dynamics. These results demonstrate the interaction of multiple topological features and dynamic states in the function of complex neural networks.

References

1. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science. **286** (1999) 509.
2. Kaiser, M., Görner, M., Hilgetag, C.C.: Criticality of spreading dynamics in hierarchical cluster networks without inhibition. New J. Phys. **9** (2007) 110.

3. Sporns, O.: Small-world connectivity, motif composition, and complexity of fractal neuronal connections. Biosystems. **85** (2006) 55.
4. Bak, P., Chen, K., Tang, C.: A forest-fire model and some thoughts on turbulence. Phys. Lett. A. **147** (1990) 297; Drossel, B., Schwabl, F.: Self-organized critical forest-fire model. Phys. Rev. Lett. **69** (1992) 1629.
5. Müller-Linow, M., Hilgetag, C.C., Hütt, M.T.: Organization of excitable dynamics in hierarchical biological networks. PLoS Comput. Biol. **4** (2008) e1000190.

Effects of Additive Gaussian Noise on Neuronal Firings in a Heterogeneous Neuronal Network

Xiaojuan Sun, Shujuan Lü, and Qishao Lu

Abstract The effects of additive Gaussian white noise on neuronal firings of a heterogeneous neuronal network are discussed in this paper. Heterogeneity is introduced through a multiplicative noise term. We show that the mean firing rate and spatiotemporal order of the studied network can attach to a maximum value at some intermediate noise intensities of the additive noise. Furthermore, we find that the optimal noise intensity needed increases with the coupling strength of the neuronal network. And we also give some comparisons with other works.

Keywords Spatiotemporal order · Mean firing rate · Neuronal network

1 Introduction

Neurons are usually subject to random fluctuations on different scales, ranging from channel noise created by random ion flow across the plasma membrane to synaptic noise created by the activity of other neurons. Based on some research results, it has been shown that noise can also play a constructive role in neuronal systems [1, 2]. Meanwhile, diversities in parameters are also very important in neuronal systems as shown in [3, 4]. It has been shown that the coherence of the studied neuronal systems can be enhanced at some intermediate diversity level, which is named as diversity-induced coherence resonance. In this paper, we will discuss the effects of additive Gaussian white noise on firing synchronization and spatiotemporal order of a neuronal network which is coupled by $N \times N$ heterogeneous Rulkov maps.

X. Sun (✉)
Zhou Pei-Yuan Center for Applied Mathematics, Tsinghua Univeristy, Beijing 100084, China; Department of Dynamics and Control, Beihang University, Beijing 100191, China
e-mail: sunxiaojuan.bj@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_34,

2 Equations of the Network

The Rulkov map [5, 6] is employed to model the dynamical behavior of neurons constituting the examined neuronal network. We consider a diffusively coupled $N \times N$ network with Rulkov map as local unit, the equations are shown as:

$$\begin{cases} u_{n+1}(i,j) = \frac{\alpha_0 + \xi_n(i,j)}{1+u_n^2(i,j)} + v_n(i,j) + D[u_n(i+1,j) + u_n(i-1,j) \\ \qquad u_n(i,j+1) + u_n(i,j-1) - 4u_n(i,j)] + \eta_n(i,j) \\ \qquad v_{n+1}(i,j) = v_n(i,j) - \beta u_n(i,j) - \gamma \end{cases} \tag{1}$$

where $u_n(i,j)$is the membrane potential of neuron $(i, j)$and $v_n(i, j)$is the corresponding ion concentration at the discrete time n. The system parameters are α_0, β and γ, whereby the latter two determine the time scale associated with the dynamics of the slow variable $v_n(i, j)$ and α_0 is the main bifurcation parameter in the absence of $\xi_n(i, j)$. If not stated otherwise, we set $\alpha_0 = 1.99$ and $\beta = \gamma = 0.001$. Each neuron is coupled diffusively with its four nearest neighbors with periodic boundary conditions given by $u(0, j) = u(N, j)$, $u(N + 1, j) = u(1, j)$, $u(i, 0) = u(i, N)$, $u(i, N+1) = u(i, 1)$. D is the coupling strength between neurons on the $N \times N$ spatial grid. Here N is taken as 128. $\xi(i, j)$ and $\eta(i, j)$ are noise terms, with the properties:

$$\begin{cases} \langle \xi_n(i,j) \rangle = 0, \\ \left\langle \xi_n(i,j)\xi_m(i^{'}, j^{'}) \right\rangle = 2\sigma_m \delta_{i,i^{'}} \delta_{j,j^{'}} \delta_{n,m} \end{cases} \tag{2}$$

and

$$\begin{cases} \langle \eta_n(i,j) \rangle = 0, \\ \left\langle \eta_n(i,j)\eta_m(i^{'}, j^{'}) \right\rangle = 2\sigma_a \delta_{i,i^{'}} \delta_{j,j^{'}} \delta_{n,m} \end{cases} \tag{3}$$

where σ_a is the noise intensity of additive Gaussian white $\eta(i, j)$. And σ_m denotes the heterogeneities of the network, the larger the $\xi_n(i, j)$ is, the more heterogeneous the network is.

3 Measurement

As stated above, we will study the effects of additive Gaussian white noise on firing synchronization and spatiotemporal order of the $N \times N$ neuronal network. Thus, we will introduce two measurements, one is the mean firing rate Π [7] and the other is the linear spatial cross-correlation measure S [8]. They can quantify the two properties, respectively. The mean firing rate is defined as

$$\Pi = \langle \pi(n) \rangle_T = \left\langle \frac{1}{N^2} \sum\nolimits_{ij} \theta[u_n(i,j) - u_{th}] \right\rangle_T \tag{4}$$

where $u_{th} = -0.2$ is the firing threshold determined by the action potential of the Rulkov neuron. Notably, $\theta(x) = 1$ if $x \geq 0$ and $\theta(x) = 0$ if $x < 0$. The bracket $\langle\rangle$ indicates the average over the whole iteration time T. The larger the Π is, the more synchrony the network is. The linear spatial cross-correlation measure S is defined as

$$S = \left\langle \frac{\mathrm{Cov}(n)}{\mathrm{Var}(n)} \right\rangle_T \tag{5}$$

where the bracket $\langle\rangle_T$ denotes averaging over the total iterated time T. Var(n) is the spatial variance at the iterated time n given as

$$\mathrm{Var}(n) = \frac{1}{N^2} \sum_{ij} (u_{i,j} - \bar{u})^2 \tag{6}$$

where $\bar{u} = N^{-2} \sum_{ij} u_{i,j}$; $Cov(n)$ is the purely spatial auto-covariance of nearest neighbors, and is defined as

$$\mathrm{Cov}(n) = \frac{1}{N^2} \sum_{ij} \frac{1}{\left|\mathbb{N}_{i,j}\right|} \sum_{b \in \mathbb{N}_{ij}} (u_{ij} - \bar{u})(b - \bar{u}) \tag{7}$$

with b consisting of all $\left|\mathbb{N}_{ij}\right| = 4$ elements of a von Neumann neighborhood $\mathbb{N}_{ij}$ at each lattice site u_{ij}. Obviously, the quantity S is efficient in analyzing nearest-neighbor relationships in space and time. Neuronal networks with larger S possess more ordered firing behavior.

4 Results

In this section, we will study the effects of additive Gaussian white on the mean firing rate Π and spatiotemporal order S of the two-dimensional heterogeneous neuronal network. Here, we take σ_m as 1×10^{-6}. From Fig. 1, we can see that there exist some intermediate noise intensities of $\eta(i, j)$, such that $\Pi$and S can reach to a maximum value. Moreover, with the coupling strength D increasing, the optimal noise intensity corresponds to maximum value of Π and S also increases. It is different from the phenomenon observed in homogeneous neuronal network [9], where they found that the optimal noise intensity is robust to the coupling strength D. Because $\alpha_0 = 1.99$, the neuronal network will be a mixture of excitable and oscillatory neurons [5, 6]. The excitable neurons inside the network will act as inhibitory ones, and its inhibitory ability will enhance with increasing of the coupling strength D. Thus, with D increasing, the neuronal network becomes harder and harder to be excited. It will result in that noise intensity needed to excite the network and the corresponding optimal noise intensity both increases with D, as shown in Fig. 1.

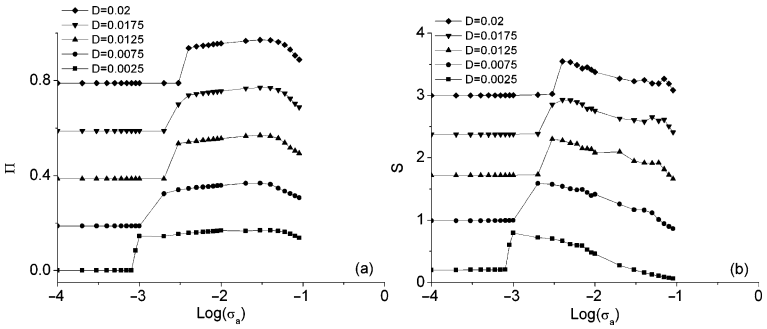


Fig. 1 Variations of the mean firing rate Π (**a**) and spatiotemporal order S (**b**) with respect to noise intensity of the additive noise $\eta(i,j)$ for various coupling strength D. Here $\sigma_m = 1 \times 10^{-6}$

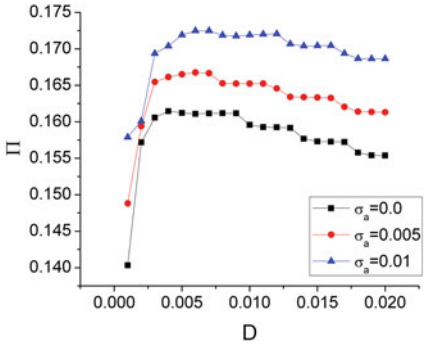


Fig. 2 Variations of the mean firing rate Π with respect to the coupling strength D for various noise intensity σ_a of the additive noise $\eta(i, j)$

Wang et al. [10] have shown that the mean firing rate of neuronal network, which are mixed by excitable and oscillatory neurons, is robust to the additive Gaussian white noise. In other words, the additive Gaussian white noise has no effects on the mean firing rate of heterogeneous neuronal network. But for the heterogeneous neuronal network discussed here, this robustness will disappear. In order to show it clearly, we take σ_m as 1×10^{-4} in Fig. 2. From this figure, we can obviously see Π increases with the noise intensity σ_a for various D. Namely, additive noise can enhance the mean firing rate of the studied heterogeneous neuronal network, which is different from the results obtained in [10]. This difference indicates that the heterogeneities introduced by noise term and uniform distribution have some inherent distinctness.

5 Conclusions

To be summarized, the effects of additive Gaussian white noise on neuronal firings in a heterogeneous neuronal network are studied in this article. The obtained numerical results have shown that there exist some intermediate noise intensities,

which can make the firing synchronization and the spatiotemporal order of the network become optimal. And the dependence of firing synchronization on additive Gaussian white noise in heterogeneous networks is different from the ones in homogeneous networks. In addition, we find that neuronal networks with different kinds of heterogeneities may show different level of firing synchronization.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Nos.10872014 and 10972018).

References

1. Pikovsky, A.S., Kurths, J.: Coherence resonance in a noise-driven excitable system. Phys. Rev. Lett. **78** (1997) 775–778.
2. Zhou, C.S., Kurths, J.: Spatiotemporal coherence resonance of phase synchronization in weakly coupled chaotic oscillators. Phys. Rev. E. **65** (2002) 040101.
3. Gassel, M., Glatt, E., Kaiser, F.: Doubly diversity-induced resonance. Phys. Rev. E. **76** (2007) 016203.
4. Chen, H.S., Zhang, J.Q., Liu, J.Q.: Enhancement of neuronal coherence by diversity in coupled Rulkov-map models. Physica A. **387**(2008) 1071–1076.
5. Rulkov, N.F.: Regularization of synchronized chaotic bursts, Phys. Rev. Lett. 86(2001) 183–186.
6. Rulkov, N.F.: Modeling of spiking-bursting neural behavior using two-dimensional map, Phys. Rev. E. **65** (2002) 041922.
7. Rieke, E., Warland, D., van Steveninck, R.R., Bialek, W.: Spikes: Exploring the Neural Code. Cambridge: MIT Press (1997) 28–38.
8. Busch, H., Kaiser, F.: Influence of spatiotemporally correlated noise on structure formation in excitable media. Phys. Rev. E. **67** (2002) 041105.
9. Xiaojuan Sun, Qishao Lu, Matjaz Perc, Jürgen Kurths: Effects of correlated Gaussian noise on the mean firing rate and correlations of an electrically coupled neuronal network, Chaos **20** (2010) 033116.
10. Wang, Q.Y., Duan, Z.S., Huang, L., et al. Pattern formation and firing synchronization in networks of map neurons. New J. Phys. **9** (2007) 383.

Stochastic Evolution Model of Neuronal Oscillator Population Under the Condition of the Variable Higher Order Coupling

Xiaodan Zhang, Rubin Wang, Zhikang Zhang, Xianfa Jiao, and Jianting Cao

Abstract In the premise of analysis on the dynamic characteristics of the transmission mechanism among the synapses, this paper modified the coupling term in the P.A. Tass's stochastic evolution model of neuronal oscillator population, introducing the variable higher order coupling term. Then, we performed the numerical simulation on the modified model. The simulation results show that without the external stimulation, the variable coupling mechanism can induce the transition between different clustering states of the neuronal oscillator population. And a full desynchronization state can exist in the middle of the transition between two different synchronization states induced by the variable coupling mechanism.

Keywords Neuronal oscillator population · Synchronization · Variable higher order coupling · Clustering · Full desynchronization

1 Introduction

Since 1989, Gary et al. [1] discovered the stimulation-induced γ synchronous oscillation (the frequency is 35–70 Hz) on the cat's primary visual cortex, the synchronous oscillation of the neuronal population attracts more and more attention in the neuroscience community. Abundant animal experiments indicated that synchronous oscillation may be the important mechanism of the information integration in the particular cortical area or between different cortical areas [1, 2]. It is approved by experiments that synchronous oscillation of the nervous activity is closely related with the cognitive function [3, 4]. In order to quantitatively describe the synchronous oscillation activity of the neuronal population, based on the fundamental characteristics of the activity of the neuron, the method that simulating the neuron as a

X. Zhang (✉)
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, China
e-mail: jimy.zhang@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_35,

periodic oscillator, simulating the neuronal oscillator population as the globally coupling neuronal oscillator network, and using the nonlinear vibration theory to study the synchronous oscillation of the nervous system and neural encoding is always been considered as a simple and effective method. In the 1960s, Winfree [5] began to apply his famous phase resetting theory to the study of physiological rhythm. Kuramoto et al. [6] advanced this theory, and began to use the concept of number density to describe the dynamical evolution of the phase oscillator population. Since the presence of the background noise in the brain, Tass [7] brought forward the stochastic phase resetting dynamical model on a basis of the former studies, and successfully applied it to the scientific fields such as biology and medicine. But for the sake of the convenience of the analysis process, the models using in the former experiments are set to be relatively simple, they only chose the lower order coupling among the neuronal oscillators to do the study. This paper modified the coupling term in the P.A. Tass's stochastic evolution model of neuronal oscillator population, introducing the variable higher order coupling term and studied the neural encoding of neuronal oscillator population on the modified model.

2 Stochastic Evolution Model of Neuronal Oscillator Population Under the Condition of the Variable Higher Order Coupling

Under the condition of stochastic noise, the dynamic equation [7] of the neuronal oscillator population which consists of N globally coupling neuronal oscillators takes the form as follow:

$$\dot{\psi}_j = \Omega + \frac{1}{N}\sum_{k=1}^{N} M(\psi_j - \psi_k) + F_j(t) \tag{1}$$

In this case, for the sake of simplicity let us assume that all oscillators have the same eigenfrequency given by Ω. The oscillators' mutual interactions are modeled by a mean-field-coupling, i.e. every oscillator is coupled to all others with equal strength $M(\psi_j - \psi_k)$ which refers to the interaction between the kth oscillator with the jth oscillator. $\psi_j - \psi_k$ is the phase difference between the two. $F_j(t)$ refers to the random forces. The coupling interaction among the neuronal oscillators takes the form:

$$M(\psi_j - \psi_k) = -\sum_{m=1}^{4} \{K_m(t)\sin[m(\psi_j - \psi_k)] + C_m(t)\cos[m(\psi_j - \psi_k)]\} \tag{2}$$

For the sake of simplicity the random forces are modeled by Gaussian white noise which is delta-correlated with zero mean value:

$$< F_j(t) >= 0\,, < F_j(t)F_k(t') >= Q\delta_{jk}\delta(t - t') \tag{3}$$

The dynamics of the Langevin Eq. (1) can be investigated by means of the corresponding Fokker-Planck equation:

$$\frac{\partial f(\{\psi_l\},t)}{\partial t}+\sum_{k=1}^{N}\frac{\partial}{\partial\psi_k}\{\frac{1}{N}\sum_{j=1}^{N}[\Omega+M(\psi_j-\psi_k)]f(\{\psi_l\},t)\}-\frac{Q}{2}\sum_{k=1}^{N}\frac{\partial^2 f(\{\psi_l\},t)}{\partial\psi_k^2}=0 \tag{4}$$

In this case, $f(\{\psi_l\},t)$ refers to probability density, where $\{\psi_l\}$ denotes the vector $(\psi_1,\cdots,\psi_N)$. $f(\{\psi_l\},t)d\psi_1\cdots d\psi_N$ provides us with the probability of finding the phases of the oscillators at time t in the intervals $\psi_j\cdots\psi_j+d\psi_j(j=1,\cdots,N)$.

Instead of analyzing the phase of every single oscillator, we restrict ourselves to the question of how many oscillators most probably have phase ψ at time t. In this way we introduce the number density $\tilde{n}$ of the oscillators which have phase ψ

$$\tilde{n}(\{\psi_l\};\psi)=\frac{1}{N}\sum_{k=1}^{N}\delta(\psi-\psi_k) \tag{5}$$

The stochastic aspect of the dynamics is taken into account by introducing the average number density n of the oscillators which have phase ψ according to

$$n(\psi,t)=<\tilde{n}\{\psi_l\};\psi>_t=\int_0^{2\pi}\cdots\int_0^{2\pi}\tilde{n}(\{\psi_l\};\psi)f(\{\psi_l\};t)d\psi_1\cdots d\psi_N \tag{6}$$

with (4), (5) and (6) one obtains

$$\frac{\partial n(\psi,t)}{\partial t}=-\frac{\partial}{\partial\psi}\{n(\psi,t)\int_0^{2\pi}M(\psi-\psi')n(\psi',t)d\psi'\}-\Omega\frac{\partial}{\partial\psi}n(\psi,t)+\frac{Q}{2}\frac{\partial^2 n(\psi,t)}{\partial\psi^2} \tag{7}$$

In order to investigate the partial differential Eq. (7) one has to take into account two boundary conditions:

$$1.\ n(0,t)=n(2\pi,t) \tag{8}$$

$$2.\int_0^{2\pi}n(\psi,t)d\psi=1 \tag{9}$$

For the convenience of observation, we adopt the firing density as $p(t)=n(0,t)$.

3 Analysis of the Numerical Calculation Results

All the simulations in this paper adopt the same frequency $\Omega=2\pi$, i.e. the period is given as $T=1$. In the Fig. 1 the coupling term is given as

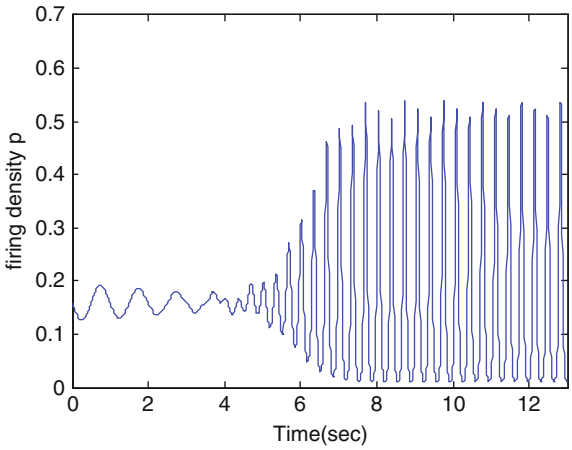


Fig. 1 The evolution of the firing density under the condition of first-order exponential decay coupling and third-order constant-coefficient coupling

$$M(\psi_j - \psi_k) = -K_1 e^{-t/\tau} \sin(\psi_j - \psi_k) - K_3 \sin[3(\psi_j - \psi_k)] \quad (\tau = 1,\ K_1 = 1,\ K_3 = 3)$$

The initial condition of the average number density is given as $n(\psi, 0) = (1/2\pi) \times (1 + 0.2 \sin \psi)$.

From Fig. 1 we can see that in the earlier period of the evolution of the neuronal oscillator population, the firing density only has one peak in one period, i.e. the neuronal oscillator population is in the one-clustering state in this process. Whereas, with the decay of the first-order coupling term, the firing density arises three peaks in one period, i.e. the neuronal oscillator population transits from the one-clustering state into the three-clustering state. The effect of the third-order constant-coefficient coupling gradually appears. What is worthy of note is that without the variable first-order coupling, only the first-order initial condition of the average number density and the third-order constant-coefficient coupling can not turn the neuronal oscillation population into the three-cluster synchronization state [8]. So we can see that although the variable coupling almost decayed to 0 in the later period of the evolution, it still played an important role in the transition of making the new synchronization.

From more simulation results we know that, the above transition not only emerges in the transition from one-cluster to three-cluster, but also emerges in transitions between other kinds of clustering state.

Figure 2 shows transition of the neuronal oscillator population from three-cluster state to one-cluster state. In the Fig. 2 the coupling terms is given as

$$M(\psi_j - \psi_k) = -K_3 e^{-t/\tau} \sin[3(\psi_j - \psi_k)] - K_1 \sin(\psi_j - \psi_k) \quad (\tau = 1,\ K_3 = 5\,K_1 = 5)$$

The initial condition of the average number density is given as

$$n(\psi, 0) = \frac{1}{2\pi} \times [1 + 0.2 \sin(3\psi)]$$

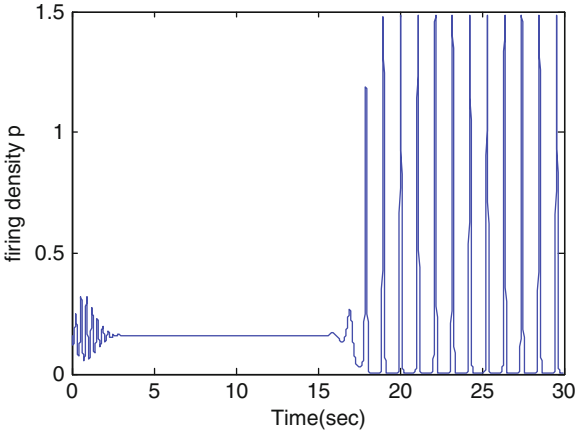


Fig. 2 The evolution of the firing density under the condition of third-order exponential decay coupling and first-order constant-coefficient coupling

In the evolution process in Fig. 2, the synchronization state induced by the exponential decay higher-order coupling will trail off according to the damping of its coupling strength. While in this process, the initial condition of average number density doesn't contain the same order term with the constant-coefficient coupling term, so the effect of the constant-coefficient coupling term doesn't exhibit in the earlier period of the evolution, but after a special time point the effect of the constant-coefficient coupling term will specially manifest. This time point is closely related with the strengths of the exponential decay higher-order coupling term and the constant-coefficient coupling term and the initial condition of the model.

From Fig. 2 we can draw another conclusion that a full desynchronization state can exist in the middle of the transition between two different synchronization states induced by the variable coupling terms. With the bating of the variable coupling term, the neuronal oscillator population can not retain the stable synchronization state and turn into a full desynchronization state. In this full desynchronization period the value of the firing density is unconverted, i.e. the neuronal oscillator population is totally scattered during this period. Whereas, after a period of full desynchronization state, the neuronal oscillator population can be turn into a new synchronization state by the effect of another coupling term which has different order with the initial condition of the average number density.

4 Conclusion

According to the numerical calculation and result analysis, we can draw the conclusions as follow:

(1) Without the external stimulation, transition of the clustering state of the neuronal oscillator population induced by variable coupling mechanism not only emerges in the transitions from the lower-order clustering state to the higher-order clustering state, but also emerges in the transitions from the higher-order

clustering state to the lower-order clustering state. Only if the neuronal oscillator population has the appropriate variable coupling structure, the neuronal oscillator population can turn from any clustering state into any another clustering state.

(2) In the neuronal oscillation population which contains the variable coupling, a full desynchronization state can exist in the middle of the transition between two different synchronization states induced by the variable coupling terms. Whereas, after a period of full desynchronization state, the neuronal oscillator population still can be turn into a new synchronization state by the effect of another coupling term which has different order with the initial condition of the average number density.

Acknowledgments Project (10872068, 10672057) supported by National Natural Science Foundation of China (NSFC)

References

1. Gray C.M., Singer, W.: Stimulus-specific neuronal oscillation in orientation columns of cat visual cortex. Proc. Natl. Acad. Sci. U. S. A. **86** (1989) 1698–1702.
2. Eckhorn, R., Bauer, R., Jordan, W., et al.: Coherent oscillations: a mechanism of feature linking in the visual cortex. Biol. Cybern. **60** (1988) 121–130.
3. Vaadia, E., Haalman, I., et al.: Dynamics of neuronal interactions in the monkey cortex in relation to behavioral events. Nature. **373** (1995) 515–518.
4. Abeles, M., Prut, Y.: Temporal structure of cortical activity. Proceedings of the International Conference, Bochum, Germany, July 16–19 (1996).
5. Winfree, A.T.: The Geometry of Biological Time. (Biomathematics. vol. 8). New York, NY, Heidelberg, Berlin: Springer-Verlag (1980).
6. Kuramoto, Y.: Chemical Oscillations, Waves and Turbulence. New York, NY: Springer (1984).
7. Tass, P.A.: Phase Resetting in Medicine and Biology. Berlin: Spring-Verlag (1999).
8. Xiaodan, Z., Runbin, W., et al.: Stochastic evolution model of neuronal oscillator population under the condition of the higher order Coupling. The 5th International Conference on Natural Computation (ICNC'09), Tianjin (2009).

Qualitative Analysis in Locally Coupled Neural Oscillator Network

Yong Meng, Yuanhua Qiao, Jun Miao, Lijuan Duan, and Faming Fang

Abstract The paper investigates a locally coupled neural oscillator autonomous system qualitatively. To obtain analytical results, we choose an approximation method and obtain the set of parameter values for which an asymptotically stable limit cycle exists, and then give sufficient conditions on the coupling parameters which can guarantee asymptotically global synchronization of oscillators given the same external input. The above results are potentially useful to analytical and numerical work on the binding problem in perceptual grouping and pattern segmentation.

Keywords Neural network · Limit cycle · Synchronization · Dynamic system

1 Introduction

A fundamental aspect of perception is to bind spatially separate sensory features to form coherent objects. There is also wide experiment evidence that perception of a single object (especially in the visual scene) involves distributed in a highly fragmented over a large spatial region. The problem thus arise of how the constituent features are correctly integrated together to represent a single object.

Some authors [1–3] assume that these features of an object are grouped based on the temporal correlation of neural activities. Thus neurons that fire in synchronization would signal features of the same object, and groups desynchronized from each other represent different objects. Experimental observations of the visual cortex of animals show that synchronization indeed exists in spatially remote columns and phase-locking can also occur between the striate cortex and extratriate cortex,

Y. Meng (✉)
College of Applied Science, Beijing University of Technology, Beijing 100124, China
e-mail: mengyong2007@emails.bjut.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_36,

between the two striate cortices of the two brain hemisphere, and across the sensorimotor cortex. These findings have concentrated the attention of many researchers on the use of neural oscillators such as Wilson–Cowan oscillators and so on. In this scheme, neural oscillators that are in phase would represent a single object (binding), while neural groups with no phase lock would represent different objects. Though there are some analysis results [3–7] on Wilson–Cowan neural network, the results on autonomous Wilson–Cowan network system are still less. To make use of oscillation in phase, it is necessary to study the autonomous Wilson–Cowan network system. In this paper, we use a neural network based on locally coupled Wilson–Cowan oscillators to analyse the binding problem. To solve the binding problem, it is necessary to determine the conditions under which neural oscillators would exhibit periodical behavior and synchronize asymptotically.

2 Mathematical Model

In the model each oscillator is described by means of simplified Wilson–Cowan equations. Such a model consists of two nonlinear ordinary differential equations representing the interactions between two populations of neurons that are distinguished by the fact that their synapses are either excitatory or inhibitory. Thus, each oscillator consists of a feedback loop between an excitatory unit x_i and inhibitory unit y_i that obey the equations:

$$\begin{aligned}\frac{dx_i}{dt} &= -r_1 x_i + r_1 H(ax_i - cy_i + I_i - \phi_x) + \alpha[\gamma_{i-1,i}(x_{i-1} - x_i) + \gamma_{i+1,i}(x_{i+1} - x_i)] \\ \frac{dy_i}{dt} &= -r_2 y_i + r_2 H(bx_i - dy_i - \phi_y) + \beta[\gamma_{i-1,i}(y_{i-1} - y_i) + \gamma_{i+1,i}(y_{i+1} - y_i)]\end{aligned} \tag{1}$$

Both x_i and y_i variables are interpreted as the proportion of active excitatory and inhibitory neurons respectively, which are supposed to be continuous variables and their values may code the information processed by these populations. Especially, the state $x_i = 0$ and $y_i = 0$ represents a background activity. The parameters have the following meanings: a is the strength of the self-excitatory connection, d is the strength of the self-inhibitory connection, b is the strength of the coupling from x to y, c is the strength of the coupling from y to x. Both ϕ_x and ϕ_y are thresholds, r_1 and r_2 modify the rate of change of the x and y unit respectively. Figure 1 shows the connections for single oscillator and the structure of an open chain of coupled oscillators. All these parameters have nonnegative values, I_i is external input to the oscillator in position i which corresponds to a pixel in object. $H(\cdot)$ is a sigmoid activation function defined as: $H(z) = 1/(1 + e^{-z/T})$, T is a parameter that sets the central slope of the sigmoid relationship, α and β represent the strength of the connection between neurons.

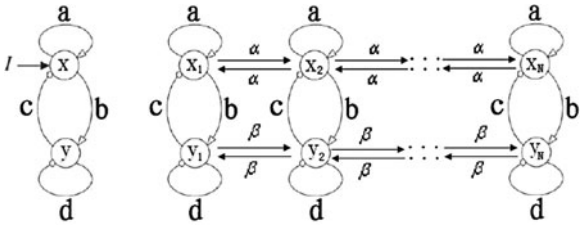


Fig. 1 A single oscillator and an open chain of coupled oscillators

3 Model analysis

3.1 Oscillating Conditions for Single Oscillator

Consider the following system of a single oscillator:

$$\begin{aligned} \tau_1 \frac{dx_i}{dt} &= -x_i + H(ax_i - cy_i + I_i - \phi_x) \\ \tau_2 \frac{dy_i}{dt} &= -y_i + H(bx_i - dy_i - \phi_y) \end{aligned} \tag{2}$$

where $\tau_1 = 1/r_1$, $\tau_2 = 1/r_2$. In order to study the equation of the model, we choose the following piece-wise linear function to approximate the sigmoid function in (2):

$$G(z) = \begin{cases} 0, & z < -2T \\ \frac{z}{4T} + \frac{1}{2}, & -2T \le z \le 2T \\ 1, & z > 2T \end{cases} \tag{3}$$

Thus, the system (2) may be described as:

$$\begin{aligned} \tau_1 \dot{x}_i &= -x_i + \eta \\ \tau_2 \dot{y}_i &= -y_i + \gamma \end{aligned} \tag{4}$$

where $\eta, \gamma \in \left\{0, 1, \frac{ax_i - cy_i + I_i + 2T - \phi_x}{4T}, \frac{bx_i - dy_i + 2T - \phi_y}{4T}\right\}$.

In order to solve the binding problem, we aim to find the conditions under which the oscillators keep silent when I_i equal to zero and the system will exist an asymptotic stable limit cycle when I_i adopts proper value. Through analysis, it is easy to find the necessary conditions for the above result.

(1)$\phi_y + 2T < b$; (2)$I_i - \phi_x - 2T > 0$; (3)$a - \phi_x + 2T + I_i < c$; (4)$d + \phi_y - 2T > 0$;
(5)$\phi_x - 2T > 0$; (6)$(a - 4T)r_1 > (d + 4T)r_2$; (7)$bc + (4T - a)(4T + d) > 0$;
(8) $|ax^* - cy^* + I_i - \phi_x| \le 2T$; (9) $\left|bx^* - dy^* - \phi_y\right| \le 2T$

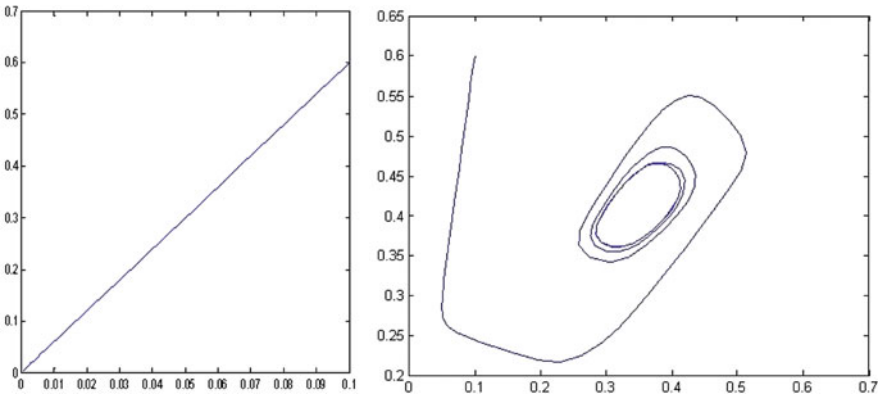


Fig. 2 Two phase diagrams (an asymptotically stable zero solution and an asymptotically stable limit cycle). The parameters are: $a = 1, b = 1, c = 2, d = 0.5, I_i = 0, I_i = 0.65$ and $T = 0.025, r_1 = r_2 = 1, \phi_x = 0.2, \phi_y = 0.15$

where

$$x^* = \frac{d+4T}{b}y^* + \frac{\phi_y - 2T}{b}; \quad y^* = \frac{b(\phi_x - I_i - 2T) - (a-4T)(\phi_y - 2T)}{(a-4T)(d+4T) - bc}.$$

On the base of the obtained conditions, it is easy to verify the fact that the above conditions satisfy Poincare–Bendixon Theorem. So an asymptotic stable limit cycle must exist. If you run the MatLab simulation, you will see that an asymptotic stable limit cycle does indeed exist and it is plotted in Fig. 2.

3.2 Synchronization of Locally Coupled Wilson–Cowan Oscillators

The boundary conditions of the coupled system (1) are as follows:

$$x_0 = x_1, \quad x_{N+1} = x_N; \quad y_0 = y_1, \quad y_{N+1} = y_N$$

$$r_{ij} = \begin{cases} 1, & \text{if } |I_i - I_j| < \phi \\ 0, & \text{otherwise} \end{cases}$$

where ϕ is a threshold. Based on the above restrictions, we will give the synchronization conditions of the coupled system in (1).

Theorem 1 *Consider an open chain coupled oscillators receiving the same input $I_i = I$, $i = 1, 2, \ldots N$ in (1), the synchronization state is asymptotically stable if the following conditions hold:* $0 < r_2 \le r_1$, $\beta \le \alpha$,

$$\alpha > \frac{(2a+c)r_1 + (b-8T)r_2}{16T(1-\cos\frac{\pi}{N})}, \quad \beta > \frac{(2a+c)r_1 + (b-8T)r_2}{16T} + \alpha\cos\frac{\pi}{N}$$

4 Conclusions

The paper presented a qualitative analysis on locally coupled autonomous Wilson–Cowan neural network, and gave the conditions under which the oscillator can oscillate and synchronize asymptotically. Note they are only sufficient conditions to achieve oscillation and synchronization. Some authors [5, 7, 8, 9] give some analysis results on chaotic synchronization in Wilson–Cowan neural network and put these into image segmentation. And some results [10] lack of the generality and effectiveness in application more or less. In contrast to them, the obtained results in the paper are more convenient to solve the binding problem. In the future research, the authors will apply the proposed model to image segmentation and test its performance.

Acknowledgments This research is partially sponsored by Natural Science Foundation of China (Nos.60673091, 60702031 and 60970087), Hi-Tech Research and Development Program of China (No.2006AA01Z122), Natural Science Foundation of Beijing (No.4072023 and 4102013), Beijing Municipal Foundation for Excellent Talents (No.20061D0501500211), Beijing Municipal Education Committee (No.KM200610005012), and National Basic Research Program of China (No.2007CB311100 and 2009CB320902). The paper is organized as follows. In Section 2, the mathematical model is described. Our main theoretical results are given in Section 3. Conclusions are given in Section 4.

References

1. Wang, D.L.: The time dimension for scene analysis. IEEE Trans. Neural Netw. **16**(6) (2005) 1401–1426.
2. Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. **12**(1) (1972) 1–24.
3. Zhao, L.: Chaotic synchronization for scene segmentation. Int. J. Mod. Phys. B. **17**(22, 23 & 24) (2003) 4387–4394.
4. Campbell, S., Wang, D.: Synchronization and desynchronization in a network of locally coupled Wilson–Cowan oscillators. IEEE Trans. Neural Netw. **7** (1996) 541–554.
5. Zhao, L.: Chaotic synchronization in 2D lattice for scene segmentation, Neurocomputing. **71** (2008) 2761–2771.
6. Zhao, L.: Chaotic synchronization in general network topology for scene segmentation. Neurocomputing. **71** (2008) 3360–3366.
7. Borisyuk, R., Kirillov, A.: Bifurcation analysis of a neural network model. Biol. Cybern., **66**(4) (1992) 319–325.
8. Wang, D.L.: Emergent synchrony in locally coupled neural oscillators. IEEE Trans. Neural Netw. **6** (1995) 941–948.
9. McAuly, A.D.: Comparing artificial and biological dynamical neural networks. Proceedings of SPIE – The International Society for Optical Engineering, vol. 6235, Signal processing, Sensor Fusion, and Target Recognition XV (2006).
10. Li, Z., Chen, G.: Global synchronization and asymptotic stability of complex dynamical networks. IEEE Circ. Syst. Soc. **53**(1) (2006) 28–33.

Determination of the Number Ratio Between Excitation Neurons and Inhibitory Neurons in Activities of Neural Networks

Yan Liu, Rubin Wang, Zhikang Zhang, Xianfa Jiao, and Jianting Cao

Abstract In this paper we found parameter regions to exhibit those different population states, called dividing zones including flat fading zone, rapid fading zone and critical zone. Based on the dividing zones we can choose the number ratio between inhibitory neurons and excitatory neurons and establish the couplings in a neural network. Our researches also show that the balance value, enabling the firing density to reach the dynamic balance, does not depend on initial conditions. In additions, the critical value is only related to the number ratio under the same initial conditions.

Keywords Inhibitory neural population · Excitatory neural population · Average number density · Critical state

1 Introduction

Since the 1960s, the theory of phase resetting was applied to the studying of physiological rhythm by Winfree [1], subsequently the theory was extended to study the synergistic effect of brain [2]. The theory was triumphantly used to neurologic disease such as Parkinson's disease in Medicine and Biology [3]. Furthermore, recent years the theory of the phase dynamics has been successfully used for studying neural networks and cognitive neurodynamics [4, 5]. The applications of stochastic phase dynamics in cognitive neurodynamics mainly emphasize on the studying of collective activities in neural oscillator population [6–8]. Furthermore, the studies on neural oscillator population lean to the studying of excitatory neural population. In our former studies, we proposed a stochastic nonlinear phase dynamic model in the presence of inhibitory neurons [9]. In this paper, we further analyze the stability on phase neural coding through computer stimulation.

Y. Liu (✉)
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, China
e-mail: woshizhu-1.2.3@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_37,

2 Dynamic Model in Spontaneous Behavior

The phase of every oscillator obeys the following phase dynamic equation:

$$\begin{aligned}\dot{\psi}_j &= \Omega_1 + \frac{1}{N}\left[\sum_{k=1}^{N_1} M_1\left(\psi_j - \psi_k\right) + \sum_{k=N_1+1}^{N} M_{21}\left(\psi_j - \psi_k\right)\right] + F_j(t)\ \mathrm{j} = 1...\mathrm{N}_1 \\ \dot{\psi}_j &= \Omega_2 + \frac{1}{N}\left[\sum_{k=1}^{N_1} M_{12}\left(\psi_j - \psi_k\right) + \sum_{k=N_1+1}^{N} M_2\left(\psi_j - \psi_k\right)\right] + F_j(t)\ \mathrm{j} = N_1 + 1...\mathrm{N}\end{aligned} \tag{1}$$

The dynamic model in spontaneous behavior is given by

$$\begin{aligned}\frac{\partial n}{\partial t} = &-\frac{\partial}{\partial\theta_1}\left\{n_1(\theta_1,t)\int_0^{2\pi} d\psi'\left[M_1\left(\theta_1-\psi'\right)n_1\left(\psi',t\right)+M_{21}\left(\theta_1-\psi'\right)n_2\left(\psi',t\right)\right]\right\} \\ &-\frac{\partial}{\partial\theta_2}\left\{n_2(\theta_2,t)\int_0^{2\pi} d\psi\cdot\left[M_{12}\left(\theta_2-\psi'\right)n_1\left(\psi',t\right)+M_2\left(\theta_2-\psi'\right)n_2\left(\psi',t\right)\right]\right\} \\ &-\left(\Omega_1\frac{\partial}{\partial\theta_1}n_1(\theta_1,t)+\Omega_2\frac{\partial}{\partial\theta_2}n_2(\theta_2,t)\right)+\frac{Q}{2}\left(\frac{\partial^2 n_1(\theta_1,t)}{\partial\theta_1^2}+\frac{\partial^2 n_2(\theta_2,t)}{\partial\theta_2^2}\right)\end{aligned} \tag{2}$$

Where $n_1(\theta_1, t)$ and $n_2(\theta_2, t)$ represent the average number of excitatory neural oscillator population and inhibitory neural oscillator population respectively. Ω_1 and Ω_2 are the eigenfrequencies of excitatory population and inhibitory population respectively. Q is the constant noise amplitude. The mutual coupling interactions are given by,

$$\begin{aligned}M_1\left(\psi_j-\psi_k\right) &= -K_1\sin\left(\psi_j-\psi_k\right)\ M_{12}\left(\psi_j-\psi_k\right) = -K_{11}\sin\left(\psi_j-\psi_k\right) \\ M_2\left(\psi_j-\psi_k\right) &= -L_1\sin\left(\psi_j-\psi_k\right)\ M_{21}\left(\psi_j-\psi_k\right) = -L_{11}\sin\left(\psi_j-\psi_k\right)\end{aligned} \tag{3}$$

Where, L_1, L_{11} are the coefficients of inhibitory coupling. K_1, K_{11} are the coefficients of excitatory coupling. $K_1 > 0$, $K_{11} > 0$, $L_1 < 0$, $L_{11} < 0$.

3 Analysis on Stability

3.1 Critical State

Figure 1 display the evolution processing of average number density of excitatory neural oscillator population. We can see that With increasing the coupling coefficients of inhibitory neural oscillators, the average number density shows a uptrend firstly, and then goes through a critical state, finally will be in downtrend. When coupling intensity of inhibitory population increases to a appropriate value, average number density will tend to a fixed value. It describes a critical case in Fig. 1b, when

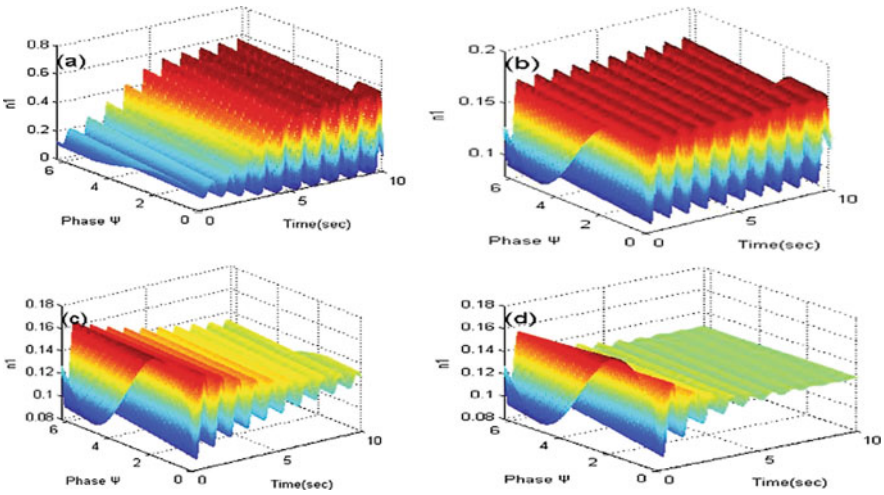


Fig. 1 Evolution of average number density with time

decreasing the inhibitory coupling, the firing density will rise and arrive dynamic balance (Fig. 1a), and when increasing the inhibitory coupling, the firing density will decline and reach to dynamic balance (Fig. 1c) or tend to a fixed value (Fig. 1d). Here, this state of Fig. 1b is defined as critical state. And the balance value under this state is named as critical value.

In this paper, the initial conditions are chose as sinusoidal form which was given by

$$n_k(\psi, 0) = \eta^* \frac{1}{2\pi} * (1 + a^* \sin(\psi)) \; \mathrm{k} = 1,2$$

$$0 < |a| \leq 1$$

Where η is a normalization factor, a is a parameter and holds. For $a = 0$, the firing density is in perfectly desynchronized state. Here, we name the critical line corresponding to $a = 0$ as totally inhibited critical line. In this paper zero was replaced by a small value 0.00001. Firing densities are approximately in uniform distribution when $a = 0.00001$.

All firing densities trend to uniform distribution beyond the totally inhibited critical line (the zoneIII) (Fig. 2). And in the zoneI the firing density rises firstly and goes into balance state. There may be the case that the firing density declines firstly and goes into balance state in the zoneII.

Figure 3 clearly shows that the size of zoneI increases with increasing excitatory coupling, but the size of zoneII is changeless. And, we can speculate that when the excitatory coupling is less than a certain value, the size of zoneI maybe reduce to zero. Because, there also is a critical value in the absence of inhibitory neurons [5].

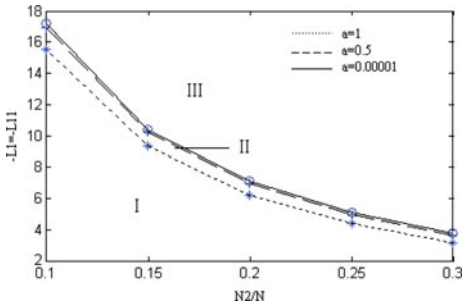


Fig. 2 Zone map when $K_1 = 3$

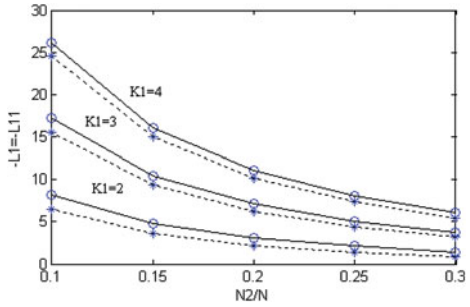


Fig. 3 Critical states in different excitatory couplings

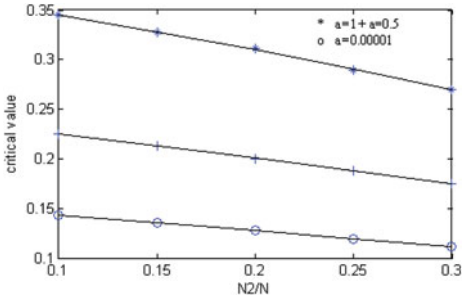


Fig. 4 Relationship between critical value and number ratio

In Fig. 4, there is only one set of curves. And this fact also conform that critical values are the same under the same initial condition. It can be seen that critical values are only related to number ratio under the same initial condition.

The equations of three lines are given by, "o" line: $y = -0.1590^{*}x + 0.1591$, "*" line: $y = -0.3750^{*}x + 0.3837$, "+" line: $y = -0.2496^{*}x + 0.2502$. If we neglect the errors, the three lines are approximately parallel. Thus the critical value and number ratio meet certain linear relationship.

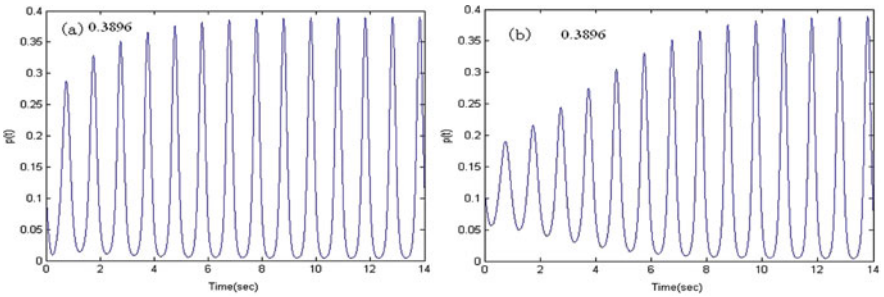


Fig. 5 Firing density (**a**) $a = 1$, (**b**) $a = 0.05$

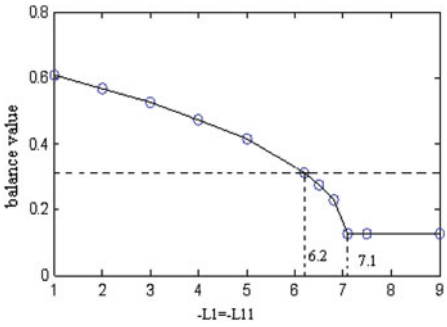


Fig. 6 Change of balance values in subarea $N_2/N = 0.2,\ K_1 = 3$

3.2 Balance Value Changes in Subarea

We can see that the critical state is related to initial conditions. We address a question of whether balance values in subarea are related to initial conditions or not.

Comparison with Fig. 5a, b, although initial conditions are different, balance values are the same. This fact shows initial conditions don't influence balance value.

The plane in Fig. 6 is divided into two parts by the dashed line. The above part corresponds to zoneI in Fig. 2. The below part correspond to zoneII and zoneIII. It can be seen that the change of balance values located in above part is flatter. But in below part balance values change rapidly. So corresponding to balance values, zoneI is named as flat fading zone. And zoneII is considered as rapid fading zone.

4 Conclusion

In this paper, we further analyze the stability on phase neural coding. And we study different states which are divided into three zones. The study on the dividing zones is considered as the reference basis of choosing the number ratio in neural network.

The results also show initial conditions don't influence balance values of firing density. And the values in critical state meet a certain linear relationship.

Acknowledgements This work is supported by National Natural Science Foundation of China (NSFC) (10872068, 10672057)

References

1. Winfree, A.T.: An integrated view of the resetting of circadian clock. Theor. Biol. **28** (1970) 327–374.
2. Haken, H.: Synopsis and Introduction, in Synergetics of the Brain. Berlin: Springer (1980).
3. Tass, P.A.: Phase Resetting in Medicine and Biology. Berlin: Springer-Verlag (1999).
4. Neiman, AB., Russell, D.F., Yakusheva, T.A., DiLullo, A, Tass, P.A.: Response clustering in transient stochastic synchronization and desynchronization of coupled neuronal bursters. Phys. Rev. E. **76** (2007) 021908.
5. Wagatsuma, H., Yamaguchi, Y.: Disambiguation of multiple sequence learning by theta phase coding. Brain Neural Netw. **12** (2005) 17–31.
6. Wang, R., Zhang, Z.: Nonlinear stochastic models of neurons activities. Neurocomputing. **51** (2003) 401–411.
7. Wang, R., Zhang, Z., Chen, G.: Energy function and energy evolution on neuronal populations. IEEE Trans. Neural Netw. **19** (2008) 535–538.
8. Wang, R., Zhang, Z.: Energy coding in biological neural network. Cogn. Neurodyn. **1**(3) (2007) 203–212.
9. Liu, Y., Wang, R., Zhang, Z., Jiao, X.: Analysis of neurodynamics on phase neural coding in the presence of inhibitory neurons. ICNC'09. Tianjin, China (2009).

Part V
Firing Oscillations and Patterns in Neuronal Networks

Transition Mechanisms Between Periodic and Chaotic Bursting Neurons

Feng Zhang, Wei Zhang, Qishao Lu, and Jianzhong Su

Abstract In this chapter, we analyze the transition mechanisms between periodic and chaotic behavior in a synaptically coupled system, which consists of bursting neurons (Hindmarsh-Rose model). We observe that chaotic bursting neurons can turn to periodic neurons or vice versa, by changes in coupling strength. The Lyaponov exponent calculations show the fine structure of a pathway to chaotic behavior when the synaptic coupling of neurons gets weaken in strength. Further we use Poincaré maps to gain more insights of the transitions.

Keywords Bifurcation · Chaos · Nonliear · Bursting neurons

1 Introduction

Neurons often exhibit bursting oscillations, as a mechanism to modulate and set pace for other brain functionalities. These bursting oscillations are distinctly characterized by a silent phase of slowly evolving steady states and an active phase of rapid firings. Based upon different bifurcation structures, one can classify bursting oscillations into several categories [1]. The mathematical analysis of bursting was studied by many authors, for example Rinzel [1], Terman [2] and Pedersen [3] more recently.

To study chaotic properties in a neuronal network, synaptically coupled bursting has also been investigated recently by several authors [4–6]. Abarbanel et al [4] has noticed the chaotic property can be regularized by coupling and that fact was further confirmed by calculating Lyaponov exponents in [4]. In an earlier paper, Su et al. [5] gave a rigorous proof that bursting solutions can be synchronized and regularized

F. Zhang (✉)
College of Mechanical Engineering, Beijing University of Technology, Beijing 100124, China
e-mail: zhangfenga@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_38,

to become periodic when the coupling strength is large enough to alter bifurcation diagram topologically.

In this paper, we start with two identical square bursting neurons that exhibit chaotic behavior, using Hindmarsh-Rose model as a example. We show that as the coupling strength increases, the chaotic behavior will be gradually replaced by synchronized regular bursting solutions. The mechanism is due to topological changes of underlying bifurcations. The bifurcation analysis and numerical evidence are in Section 2. In Section 3, we present some preliminary analysis on the transition through study of flow-induced Poincaré Maps, which provide fine detail of the dynamics structure.

2 Dynamics Behavior of Bursting Neurons

We study first the numerical solutions of coupled HR model:

$$\begin{cases} x_1{}' = z_1 + \phi(x_1) - y_1 + I - \alpha(x_1 + 1.4)H(x_2 + 0.85), \\ z_1{}' = \psi(x_1) - z_1, \\ y_1{}' = -\epsilon y_1 + \epsilon S(x_1 - c), \\ x_2{}' = z_2 + \phi(x_2) - y_2 + I - \alpha(x_2 + 1.4)H(x_1 + 0.85), \\ z_2{}' = \psi(x_2) - z_2, \\ y_2{}' = -\epsilon y_2 + \epsilon S(x_2 - c). \end{cases} \tag{1}$$

where $I = 3.281, \epsilon = 0.0021, c = -1.6, S = 4.0, \alpha = 0.2$, $\phi(s) = 3.0s^2 - s^3$, $\psi(s) = 1.0 - 5.0s^2$, the initial values are set at $x_1 = 0$, $z_1 = 0$, $y_1 = -2$ and $x_2 = 0$, $z_2 = 0.2$, $y_2 = -3.02$. The parameter $\alpha > 0$ represents the strength of the coupling, and we restrict our discussion to excitatory coupling. The function $H(\cdot)$ is the Heaviside function used by Somers et al. [7] to model the synapse action. Although the discontinuous Heaviside function is used, systems with smoothed version of H yield to similar results. In Eq. (1), when x_1 and x_2 are below -0.85, two neurons are not effected by synapses at all.

The sub-system containing the two equations of x'_i and z'_i are called fast system (FS). The equations on y'_i are called slow equations. Through numerical verification via XPP [8], we observe that the system has the following property(Fig. 1a): the set of steady states of (FS) consists of an S-shaped curve of y in (x, y)-plan denoted by S. P_λ and P_ρ are fold bifurcation points. We denote the upper, middle and lower branches of S by U, M, L. Two homoclinic points on middle branch are at $y = y_h$ and $y = y_H$.

When the coupling strength α is small, two neurons have chaotic oscillations. With parameter α increasing, we find that the solutions will quickly synchronize into a bursting solution that is regular and attracting, as shown in Fig. 1b.

The numerical calculation of bifurcations of fast system of Eq. (1) indicated that when $\alpha > \alpha_0^-$, the family of periodic solutions P_1 and P_2 will be detached from the middle branch and will merge into a continuous branch P_α. The periodic family P_α will start from one Hopf bifurcation and terminate at another Hopf-bifurcation at

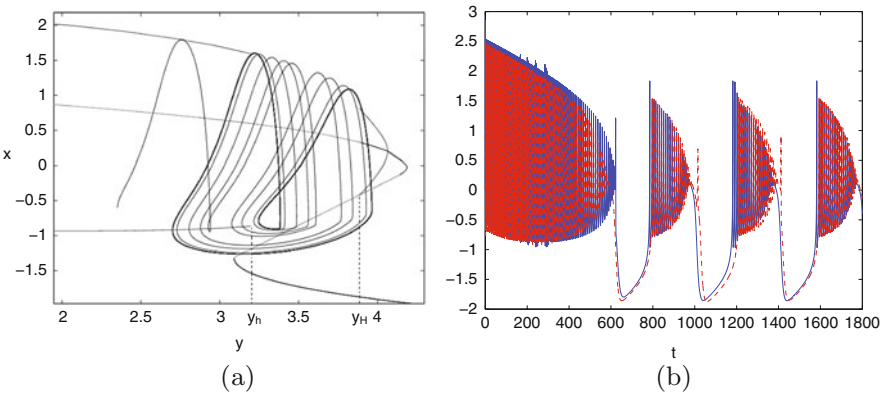


Fig. 1 (**a**) Illustration of bifurcation and phase portrait of a chaotic neuron. The dynamics near homoclinic points on the middle branch is the main reason for chaotic behavior. (**b**) Two neurons solutions in Eq. (1) quickly converge to a periodic regular bursting solution when the coupling is large enough to alter the bifurcation

upper branch at $y = y_b$ and at $y = y_B$. This finding was summerized in a rigorously proven result in [5].

3 Poincaré Maps

The remaining open question is the transition process from periodic bursting to chaotic bursting. This is the main focus for this current paper. We will address the issue through study of the flow-induced Poincaré maps and understand its route to chaos.

This chaotic square bursting was mathematically studied by Terman [2]. The chaotic trajectory are characterized by varying spike number (winding number) during each burst, for complete understanding of its Smale horses how formation, we refer the readers to two papers [2, 3].

From first Lyapunov exponent bifurcation diagram of Eq. (1) (shown in Fig. 2a), we obtain the transition from chaos to regular as α increases. Once the parameter $\alpha > \alpha_0$, two periodic branches connect to become one originated from two Hopf bifurcation points, chaotic character vanishes completely.

We next define a rectangular neighborhood $\mathbf{B_h} = \{(x, z, y) : |x| \leq \epsilon, |z| \leq \epsilon, y_1 \leq y \leq y_2\}$. where we take $y_1 = y_h - K\epsilon$ and $y_2 = y_H + K\epsilon$ and covers of the middle branch $\{(0, 0, y), y_1 \leq y \leq y_2\}$ between the homoclinic points between two homoclinic orbits.

We define a Poincaré section as

$$\Sigma_0 = \Sigma_{BK}\{(x, z, y) \in \mathbf{B_h} : x = \epsilon\}.$$

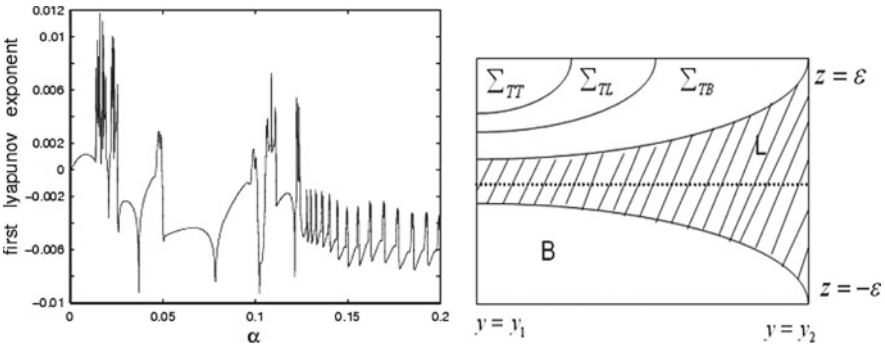


Fig. 2 (**a**) The first lyapunov exponent bifurcation diagrem of the full system (Eq. 1). (**b**) The Poincaré section Σ_0 is divided into regions of different winding numbers. Any two bursting trajectories will cross Σ_0 in such a small distance that they cannot be located in regions with a winding number difference lager than 2

Through analysis of both fast and slow manifold, we obtain the Poincaré map :

$$\begin{aligned} &p : \Sigma_0 \longrightarrow \Sigma_0 \\ &p(\epsilon, z, y) = (\epsilon, D, y + m\epsilon + \epsilon ln(\epsilon/z)/\lambda_2). \end{aligned} \tag{2}$$

where $D = a\epsilon(z/\epsilon)^{\lambda_1/\lambda_2} - b(y + \epsilon ln(\epsilon/z)/\lambda_2 - y_1)(y_2 - y - \epsilon ln(\epsilon/z)/\lambda_2)$.

The family of middle branch with slow variable y as parameter, which have one positive eigenvalue and one negative eigenvalue $-\lambda_1$ and λ_2. Using the Poincaré map we constructed, we now summarize our theoretical finding in our key theorem of this paper. We show the narrowing of the gap between two homoclinic points y_h and y_H will force the vanishing of Smale Horseshoe by restricting the possible winding number (or spike number) during each burst episode. Our results as well as a complete proof are omitted here and will be submitted in a separate paper.

Theorem 1 *We assume in equation (1) all parameters are fixed except for* ϵ *and* α. *There exists* $\epsilon_0 > 0$ *such that the fast-slow analysis becomes valid for the system if* $\epsilon < \epsilon_0$. *There exist* $\delta > 0$ *such that When* $y_H - y_h < \delta$, *any two trajectories entering* B_h *will have the difference of winding numbers* ≤ 2, *i.e., there will be no Smale Horseshoe.*

4 Conclusions and Discussion

We are interested in dynamic patterns that arise during the transaction from unsynchronized chaotic actions of neurons to ultimately synchronized regular neuronal activities as we increase the strength of synaptic coupling. We observe the dynamics near two homoclinic points played important roles at the transition.

When the slow nullcline intersects the middle branch between two homoclinic points [6], there is a different mechanism to generate chaos. We have also studied the mechanism of the transitions using a similar method.

Acknowledgments JS is partially supported by NIH grant No. 1R21NS052850-01A1. FZ and QL are supported by the National Science Foundation of China No. 10872014.

References

1. Rinzel, J.: A formal classification of bursting mechanisms in excitable systems. Proc. Int. Congr. Math. (1987) 1578–1593.
2. Terman, D.: Chaotic spikes arising from a model of bursting in excitable membranes. J. Appl. Math. **51**(5) (1991) 1418–1450.
3. Pedersen, M., Sorensen, M.: The effect of noise on beta-cell burst period. SIAM. J. Appl. Math. **67** (2007) 530–542.
4. Abarbanel, H.D.I., Huerta, R., Rabinovich, M.I., Rulkov, N.F., Rowat, P.F., Selverston, AI.: Synchronized action of synaptically coupled chaotic model neurons. Neural. Comput. **8**(8) (1996) 1567–1602.
5. Su, J., Perez, H., He, M.: Regular bursting emerging from coupled chaotic neurons. Discrete. Contin. Dyn. Syst. supplemental issue. **3**(3) (2008) 208–219.
6. Belykh, V.N., Belykh, I.V., Colding-J_rgensen, M., Mosekilde, E.: Homoclinic bifurcations leading to the emergence of bursting oscillations in cell models. Eur. Phys. J. E. **3**(3) (2000) 205–219.
7. Somers, D., Kopell, N.: Rapid synchronization through fast threshold modulation. Biol. Cybern. **68** (1993) 393–407.
8. Ermentrout, B.: Simulating, Analyzing, and Animating Dynamical Systems. Philadelphia, PA: SIAM (2002).

Dynamical Characteristics of the Fractional-Order FitzHugh-Nagumo Model Neuron

Yong Liu, Yong Xie, Yanmei Kang, Ning Tan, Jun Jiang, and Jian-Xue Xu

Abstract Through the research on the fractional-order FitzHugh-Nagumo model, it is found that the Hopf bifurcation point in such a model, where the state of the model neuron changes from the quiescence into periodic spiking, is different from that of the corresponding integer-order model when the externally applied current is considered to be the bifurcation parameter. Moreover, we demonstrate that the range of periodic spiking of the fractional-order model neuron is clearly smaller than that of the corresponding integer-order model neuron, that is, the range of periodic spiking of the former is just embedded in that of the latter. In addition, we show that the firing frequency of the fractional-order model neuron is evidently larger than that of the integer-order counterpart. The Adomian decomposition method is used to calculate fractional-order differential equations numerically due to its rapid convergence and high accuracy.

Keywords Fractional-order · Hopf bifurcation · FitzHugh-Nagumo model · Firing frequency

1 Introduction

In the 1950s Hodgkin and Huxley did their landmark work in electrophysiology and proposed the famous Hodgkin-Huxley (HH) model, which is the first mathematical model that describes how action potentials in neurons are initiated and propagated [1]. The HH model makes us possible to study the neuronal behavior from the viewpoint of dynamical systems theory. Unexpectedly, the HH theory and its model can not describe the rapid initiation and variable onset potential in cortical neurons [2]. The HH model, therefore, is not perfect. Actually it is demonstrated in [3] that the electrical properties of nerve cell membranes and the propagation of

Y. Liu (✉)
MOE Key Laboratory for Strength and Vibration, School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China
e-mail: ylhaoam@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_39,

electrical signals are well characterized by differential equations of fractional order. Brainstem vestibule oculomotor neurons, for example, exhibit the fractional-order dynamics, and the physiological basis of fractional differentiation is analyzed in detail [4]. Besides, it was reported most recently that neocortical pyramidal neurons show fractional differentiation in Nature Neuroscience [5], where the author finds that single rat neocortical pyramidal neurons adapt with a time scale that depends on the time scale of changes in stimulus statistics. This multiple time scale adaptation is consistent with fractional differentiation, such that the neuron's firing rate is a fractional derivative of slowly varying stimulus parameters. Therefore it seems that fractional-order differential equations are more suitable for description of the electrical properties of certain nerve cell membranes. Now the research on the dynamical behavior of neurons modeled by integer-order differential equations can be found everywhere, to the best of our knowledge, however, the corresponding dynamical characteristics of fractional-order model neurons are not discussed yet. Consequently, we devote ourselves to exploring the dynamical characteristics of the fractional-order FitzHugh-Nagumo (FHN) neuron in this paper.

This paper is organized as follows. Section 2 is dedicated to describing the bifurcation behavior of the Fractional-Order FHN model neuron. In Section 3 we show that the firing frequency of the fractional-order model neuron is larger than that of the integer-order counterpart. Conclusions are drawn in Section 4.

2 Bifurcation Behavior of the Fractional-Order Model Neuron

The fractional-order FHN model reads as follows:

$$\begin{cases} D_*^q x = x - \frac{x^3}{3} - y + I, \\ D_*^q y = r(x + a - by). \end{cases} \quad (1)$$

Here x denotes the membranes voltage and the recovery variable y represents activation of an outward current. Parameter I is the injected current. D_*^q is a fractional derivative operator in the sense of the Caputo definition, where q is the order of the fractional derivative. a, b and r are parameters. All of the variables in the model are dimensionless.

For $q = 1$, the model degenerates into the classical FHN model, which is investigated in detail in [6]. When $r = 1/13$, $a = 0.7$ and $b = 0.8$, the neuron shows type-II excitability if the injected current I is regarded as the control parameter. Specifically, the neuron undergoes Subcritical Adronov-Hopf bifurcation at $I = 0.3297$ where the state of the neuron changes from quiescence into periodic spiking. At $I = 1.4203$ the neuron goes through Subcritical Adronov-Hopf bifurcation again, and the state of the neuron changes from periodic spiking into quiescence.

Let (x^*, y^*) be the equilibrium point of System (1), then the linear topological equivalence of System (1) in the sufficient small neighborhood of (x^*, y^*) has the following form:

$$\begin{cases} D_*^q x = (1 - x^{*2})x - y, \\ D_*^q y = rx - rby. \end{cases} \tag{2}$$

The eigenvalues of the Jacobian matrix are:

$$\lambda_{1,2} = \frac{1 - x^{*2} - br \pm \sqrt{(x^{*2} + br - 1)^2 - 4r[1 - b(1 - x^{*2})]}}{2}. \tag{3}$$

We can obtain the condition of stability of the equilibrium point in System (1) from [7]:

1. If $|\arg(\lambda)| > \frac{q\pi}{2}$, then equilibrium point (x^*, y^*) is stable.
2. If $|\arg(\lambda)| < \frac{q\pi}{2}$, then equilibrium point (x^*, y^*) is unstable.

Under a set of parameter values, namely, $r = 1/13$, $a = 0.7$, $b = 0.8$ and $I = 0.37$, the equilibrium point of System (1) is $(-0.9339, -0.2924)$. The corresponding eigenvalues of the linearization system (2) are $\lambda_{1,2} = 0.0331 \pm 0.2607i$ in which $\left|\arg(\lambda_{1,2})\right| = 1.4445$ in radian. Let $\left|\arg(\lambda_{1,2})\right| = \frac{q_{cr}\pi}{2}$, then we can get the critical fractional order $q_{cr} = 0.9196$. If $q > q_{\text{cr}}$, the equilibrium point $(-0.9339, -0.2924)$ is unstable, and the neuron is under the state of periodic spiking. If $q < q_{cr}$, the equilibrium point is stable, and the state of the neuron is quiescent. Based on the stability criteria mentioned above, we can conclude that:

1. For $q = 1$, the equilibrium $(-0.9339, -0.2924)$ is unstable because of $q > q_{\text{cr}} = 0.9196$, and thus the neuron exhibits periodic spiking.
2. For $q = 0.5$, since $q < q_{cr} = 0.9196$, the equilibrium $(-0.9339, -0.2924)$ is stable, and the neuron is quiescent.

Thus it can be seen that under the set of parameter values of $r = 1/13$, $a = 0.7$, $b = 0.8$ and $I = 0.37$, the integer-order FHN model neuron exhibits periodic spiking while the fractional-order counterpart of order 0.5 is quiescent. Therefore the strength of the injected current making the integer-order model neuron begin to fire periodically is smaller than that making the fractional-order model neuron begin to fire periodically.

In the same way, we can check that the injected current makes the model neuron terminate periodic spiking for the two cases of the integer-order and the fractional-order, respectively, and find that the injected current in the integer-order model neuron is larger than the corresponding current in the fractional-order counterpart.

To demonstrate the relationship of the range of the injected current, where the model neuron exhibits periodic spiking, between the two cases of the integer-order and the fractional-order, we calculate their size numerically. Let $x_{\max}$ and $x_{\min}$ be the maximum and the minimum of the membrane voltage respectively when the model neuron displays periodic spiking. Figure 1 shows the curves of $x_{\max} - x_{\min}$ as the strength of injected current changes from 0 to 1.5. The solid line denotes the curve

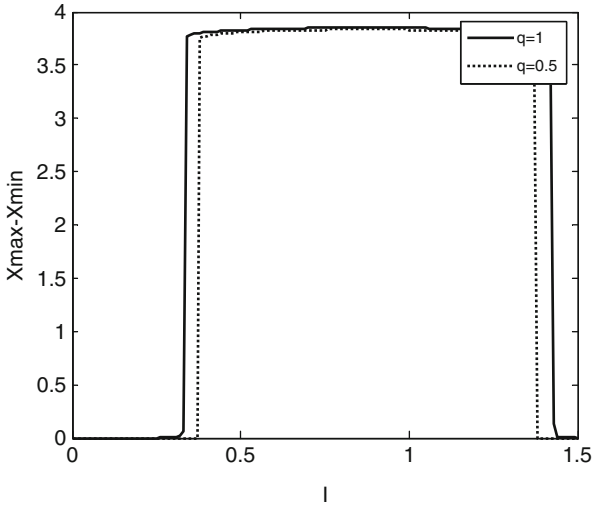


Fig. 1 Curve of $x_{\max} - x_{\min}$ as the strength of injected current changes from 0 to 1.5

of the integer-order model neuron, and the dotted line is the curve of the fractional-order counterpart. We can find that the range of the strength of the injected current making the model neuron exhibit periodic spiking in the case of the fractional-order is clearly narrower than that in the case of the integer-order from the Fig. 1. In other words, the range of the injected current of the fractional-order model neuron is embedded in the corresponding range of the integer-order model neuron. This surprising phenomenon can be explained by the stability of the equilibrium point. As we know, the state of the FHN model neuron is determined by the stability of the equilibrium point, namely, if the equilibrium is stable, then the neuron is quiescent; otherwise, the neuron exhibits periodic spiking. Furthermore, we can learn from [7] that the region of stability of the equilibrium point is extended for linear fractional-order systems, and is evidently wider than the corresponding region for linear integer-order systems. As a result, the region of instability becomes narrow, which causes the range of the strength of the injected current with periodic spiking in the fractional-order model neuron becomes smaller than that in the corresponding integer-order model neuron.

3 Firing Frequency of the Fractional-Order Model Neuron

Although the range of the strength of the injected current in the fractional-order model neuron, which can make the model neuron exhibit periodic spiking, is smaller than that in the corresponding integer-order model neuron, the firing frequency of

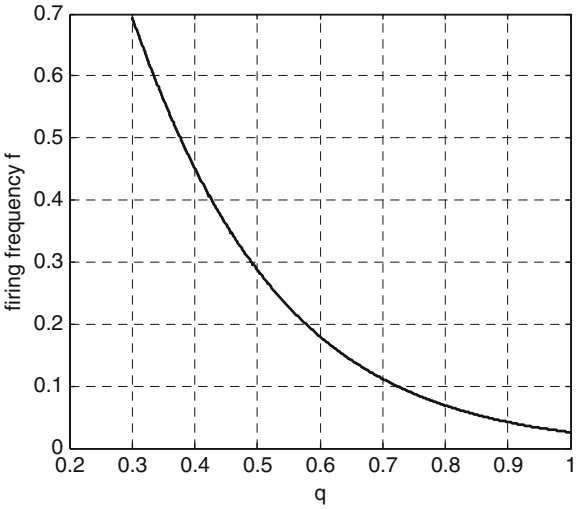


Fig. 2 The curve of the firing frequency versus the fractional order of the model neuron

the fractional-order model neuron is clearly higher than that of the integer-order counterpart for a certain injected current. Under a set of parameter values of $r = 1/13$, $a = 0.7$, $b = 0.8$ and $I = 0.6$, the curve of the firing frequency f versus the fractional order q shows adequately the properties of the firing frequency of the model neuron, as shown in Fig. 2.

4 Conclusion

As a generalization of the integer-order calculus, the fractional calculus provides a more powerful tool for the science and engineering. The fractional calculus has the potential to accomplish what integer-order calculus can not do. It is widely considered that a good many of the great future developments come from the application calculus in different fields. In this study, the bifurcation behavior of the fractional-order model neuron has been investigated. It is found that the range of the strength of the injected current that makes the fractional-order FHN model neuron fire periodically is embedded into the corresponding range where the integer-order counterpart exhibits periodic spiking. Furthermore, we have studied the firing frequency of the fractional-order model neuron. Additionally, the Adomian decomposition method has been used to calculate fractional-order differential equations numerically because of its rapid convergence and high accuracy.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No. 10502039, 10602041 and 10602003).

References

1. Hodgkin, A., Huxley, A.: A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. **117** (1952) 500–544.
2. Naundorf, B., Wolf, F., Volgushev, M.: Unique features of action potential initiation in cortical neurons. Nature. **440** (2006) 1060–1063.
3. Magin, R.: Fractional Calculus in Bioengineering. Redding, CT: Begell House, Inc. (2004).
4. Auastasio, T.: The fractional-order dynamics of brainstem vestibulo-oculomotor neurons. Biol. Cybern. **72** (1994) 69–79.
5. Lundstrom, B., Higgs, M., Spain, W., Fairhall, A.: Fractional differentiation by neocortical pyramidal neurons. Nat Neurosci. **11** (2008) 1335–1342.
6. Zhang, G., Xu, J.: Characteristic of nonlinear system stochastic resonance in the neighbourhood of bifurcation point. Acta Phys. Sin. **54** (2005) 557–564.
7. Deng, W.: Theoretical analysis and numerical computation for fractional differential equations. Ph.D dissertation, Shanghai University (2007).

Dynamics Analysis of the Hyperpolarized Model for Bursting Calcium Oscillations in Non-excitable Cells

Pan Meng and Qishao Lu

Abstract In this paper the hyperpolarized model for bursting calcium oscillations in non-excitable cells is considered. Compared with the model without hyperpolarizing current, different bursting types of oscillation are presented. Fast-slow analysis is used to study the mechanism of oscillations. The results are instructive for further understanding the dynamical behavior and mechanism of complete intracellular and intercellular calcium signaling and information processing in biological nervous systems.

Keywords Non-excitable cell · Calcium oscillation · Bursting

Bursting is a relatively slow rhythmic alternation between an active phase of rapid spiking and a quiescent phase without spiking. It occurs in many nerve and endocrine cells, including thalamic neurons, hypothalamic cells, pyramidal neurons etc. Recently, bursting oscillations were also observed during the intra-cellular calcium signaling processing. In excitable and non-excitable cells, a significant part of signal transduction, controlling the complex behavior of biological systems, is performed by the oscillatory changing of free cytosolic calcium oscillation. Borghans et al. proposed calcium-induced calcium release based on model focusing on the effect of cytosolic calcium on the degradation of inositol trisphosphate [1]. There are three variables in the model, that is, the free calcium concentration in the cytosol ($C_{ca_{\mathrm{cyt}}}$), the free calcium concentration in the endoplasmic reticulum ($C_{ca_{\mathrm{er}}}$), and the inositol trisphosphate concentration in the cytosol (C_{IP_3}). A few types of bursting calcium oscillations have been reported in this model, including simple periodic oscillation, quasi-periodic oscillation, point-cycle bursting of subHopf–subHopf type, and pint-point bursting of subHopf–subHopf type [2]. Now we add a hyperpolarizinig current I in this model and see how the dynamical behavior changes.

P. Meng (✉)
Department of Dynamics and Control, Beihang University, Beijing 100191, China
e-mail: mengpan2010@yahoo.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_40,

Consider the following hyperpolarized-model:

$$\frac{dc_{ca_{\text{cyt}}}}{dt} = J_{\text{in}} + J_{\text{leak}} - J_{\text{pump}} + J_{\text{er}} - J_{\text{out}} + I, \tag{1}$$

$$\frac{dC_{IP_3}}{dt} = r \cdot k_p - J_D - J_C, \tag{2}$$

$$\frac{dc_{ca_{\text{er}}}}{dt} = J_{\text{pump}} - J_{\text{leak}} - J_{\text{er}} \tag{3}$$

where

$$J_{\text{in}} = k_{\text{in1}} \cdot r + k_{\text{in2}},$$

$$J_{\text{leak}} = r \cdot k_{\text{leak}} \cdot \frac{C_{IP_3}^4}{C_{IP_3}^4 + K_a^4} \cdot \frac{C_{ca_{\text{er}}}^2}{C_{ca_{\text{er}}}^2 + K_y^2} \cdot \frac{C_{ca_{\text{cyt}}}^4}{C_{ca_{\text{cyt}}}^4 + K_z^4},$$

$$J_{\text{pump}} = k_{\text{pump}} \cdot \frac{C_{ca_{\text{cyt}}}^2}{C_{ca_{\text{cyt}}}^2 + K_z^2},$$

$$J_{\text{out}} = k_{\text{out}} \cdot C_{ca_{\text{cyt}}}, \quad J_{\text{er}} = k_{\text{f}} \cdot C_{ca_{\text{er}}}, \quad J_A = r \cdot k_p,$$

$$J_D = k_d \cdot \frac{C_{IP_3}^2}{C_{IP_3}^2 + K_p^2} \frac{C_{ca_{\text{cyt}}}^4}{C_{ca_{\text{cyt}}}^4 + K_z^4},$$

$$J_C = \varepsilon \cdot C_{IP_3}.$$

We choose $I = -0.3$, and other parameters used here can be found in [2]. In what follows, we study this model by hanging the parameter r, similar to [3]. Eqs. (1)–(3) is a system with both slow and fast scales [4]. The fast subsystem is composed of (1) and (2), in which the slow variable $C_{ca_{\text{er}}}$ is considered as a control parameter. The bifurcation diagrams of the fast subsystem and the phase trajectories of the whole system for different values of r are plotted in Fig. 1. Fast-slow dynamics analysis is used to study the generation mechanism of bursting oscillations of the model.

When $r = 1.08$, there is a quasi-periodic bursting due to a supercritical Neimark–Sacker (NS) bifurcation in the whole system, which is similar to that of the original model at $r = 0.99$, but there is a difference between them. In the former case, the torus does not surround the subHopf bifurcation point (see Fig. 1b), but the latter case does (see Fig. 1d).

When $r = 1.081$, there is a point-cycle bursting. Its main character is that the active and silent phases of bursting depend on a stable steady state (that is, a stable focus point) and a stable limit cycle. From Fig. 2b it is seen that the trajectory of the whole system only surround the fold limit cycles bifurcation point and there are five spikes per burst.

When $r = 1.2$, there is a point–point bursting of fold-subHopf type. The main character of point–point bursting is that the silent as well as the active phases of bursting is associated to a steady state in the fast subsystem. From Fig. 3b it is seen that the silent phase disappears due to the fold bifurcation, and the active phase

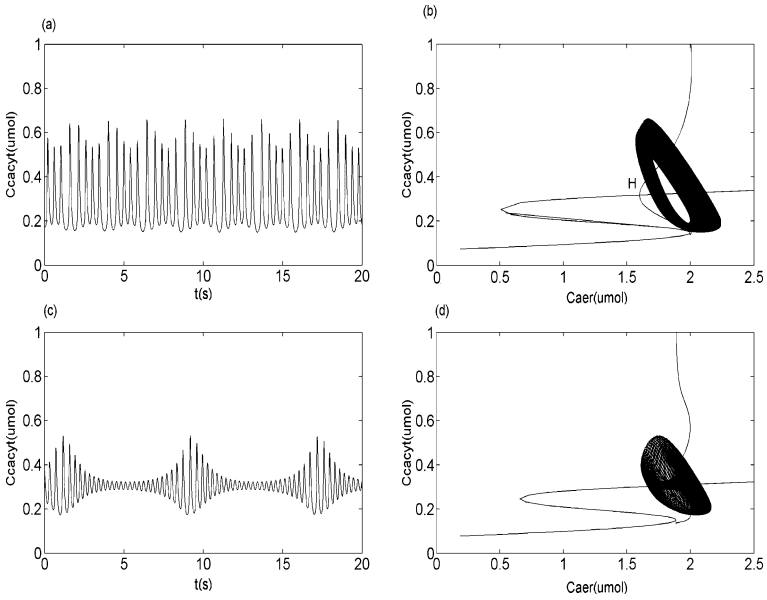


Fig. 1 **a** Time series of the hyperpolarized-model with $r = 1.08$, **b** the bifurcation diagram of the fast subsystem of the hyperpolarized-model ,where H refers to the subHopf bifurcation point, **c** time series of the original model with $r = 0.99$ ([3]), **d** the bifurcation diagram of the fast-subsystem of the original model

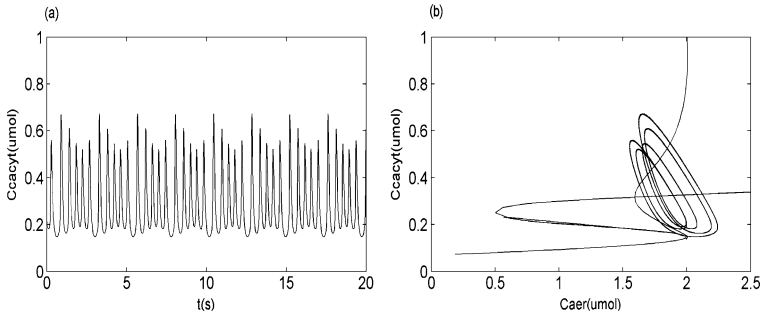


Fig. 2 **a** Time series of the hyperpolarized-model with $r = 1.081$, **b** the bifurcation diagram of the fast subsystem

ends at the subcritical Hopf bifurcation. Moreover, the slow passage effect is not so apparent as that in the original model.

When $r = 1.3$, there is a point-point bursting of subHopf–subHopf type in the hyperpolarized model. This phenomenon is similar to the original model when $r = 1.13$. As comparison, the saddle homoclinic bifurcation of the fast subsystem is far away from the fold bifurcation points in the hyperpolarized model (see Fig. 4b),

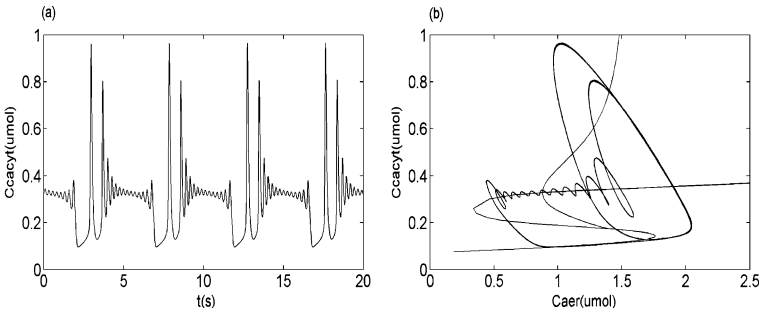


Fig. 3 **a** Time series of the hyperpolarized-model with $r = 1.2$, **b** the bifurcation diagram of the fast subsystem

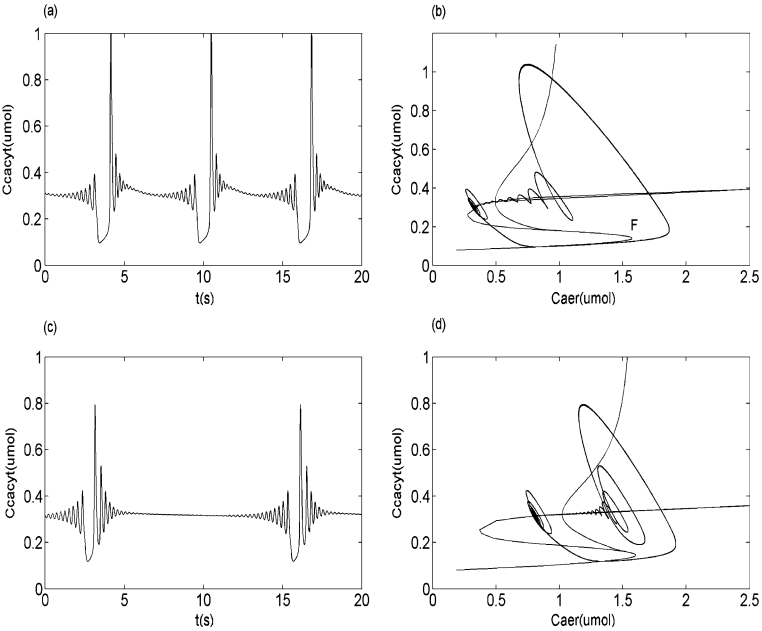


Fig. 4 **a** Time series of the hyperpolarized-model with $r = 1.3$, **b** the bifurcation diagrams of the fast subsystem of the hyperpolarized-model , where F is the fold bifurcation, **c** time series of the original model with $r = 1.13$ ([3]), **d** the bifurcation diagrams of the fast-subsystem of the original model

but the saddle homoclinic bifurcation arises when the fold bifurcation collides with the unstable periodic branch originating from the subHopf bifurcation in the origin model (see Fig. 4d). Another difference is the slow passage effect is also not so apparent as that in the original model.

Acknowledgements This work was supported by the National Natural Science Foundation of China under Grant No. 10872014.

References

1. Borghans, J.A.M., Dupont, G., Goldbeter, A.: Complex intracellular oscillations. A theoretical exploration of possible. BioSystems. **57** (2000) 75–86.
2. Perc, M., Marhl, M.: Different types of bursting calcium oscillations in non-excitable cells. Chaos Solitons Fractals. **18** (2003) 759–773.
3. Feng, Z. Qi-Shao, L., Li-Xia, D.: Dynamical analysis and transition mechanism of bursting calcium oscillations in non-excitable cells. Chin. Phys. Lett. **24** (2007) 3344–3346.
4. Rinzel, J.: Bursting oscillation in an excitable membrane model. In Lecture Notes in Mathematics, vol. 15. Berlin: Springer-Verlag (1985).

Effect of Temperature on Synchronization Phenomena in Coupled Spiking Neuron Models

Yasuomi D. Sato

Abstract Thermal effects on synchronization in a pair system of the Morris-Lecar type with excitatory synaptic couplings is studied. In the system, two parameters of temperature on ionic mechanisms (μ) and a relaxation rate of the synaptic couplings (α) are introduced. (α, μ)-modulated synchronization transition are systematically investigated by their stability analysis derived from a phase reduction method. A phase response curve (PRC) is also obtained as one of important spiking properties. The PRC is very temperature sensitive. In addition, The μ-dependent duration and firing frequency of the spikes are founded.

1 Introduction

Temperature is one of the most significant physical variables in neuronal dynamics. In particular, neuronal activities as representatives of oscillatory or nonlinear dynamics are very sensitive for temperature on ionic channel mechanisms. Thermally sensitive neurons, which mimics all spike train patterns observed in electroreceptors of dogfish, exhibits various bursting behaviors as the temperature is varied [1].

The past experiments using pacemaker showed that temperature increases cause a regular and reproducible increase in the firing frequency [2], and remarkable increase of the duration of rising and falling phases of the action potential [3]. Cocherová studied temperature relationships of the modified Hodgkin-Huxley model to the duration of an action potential [4]. However, it is not well-known about synchronization transition influenced by a temperature-modulated firing property

Y.D. Sato (✉)
Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu 808-0196, Japan; Frankfurt Institute for Advanced Studies (FIAS), Johann Wolfgang Goethe University, 60438 Frankfurt am Main, Germany
e-mail: sato-y@brain.kyutech.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_41,

measure such as a firing rate. To our knowledge, Prior and Grega showed increased synchrony of activity between two bursting neurons at low temperature [5].

A phase response curve (PRC), which is regarded as another important firing property measure, describes the transient change in the cycle period induced by a small stimulus [6]. The PRC can be classified into two neuron types, type-I and type-II. Two excitatory type-I neurons become synchronized more smoothly than the type-II neurons [7]. But, effects of temperature on the PRC, neuron type and its related synchronous behavior are still unclear. Effects of temperature on, firing duration, PRCs and synchrony in a pair of neuron models via excitatory chemical synapses are addressed.

2 A Pair of Coupled Neurons

The paired system is based on the Morris-Lecar (ML) model [8], while the chemical synapses are modeling α-function [9]:

$$C_m \frac{dV_i}{dt} = -g_L(V_i - V_L) - g_{Ca} m_\infty(V_i)(V_i - V_{Ca}) - g_K W_i(V_i - V_K) + I_{app} + \epsilon s_{\bar{i}}, \tag{1}$$

$$\frac{dW_i}{dt} = \mu \{w_\infty(V_i) - W_i\} \tau_w(V_i), \tag{2}$$

$$\frac{ds_i}{dt} = \mu\alpha(-s_i + h_i), \tag{3}$$

$$\frac{dh_i}{dt} = \mu\alpha(-h_i + H(V_i)), \quad H(V_i) = \begin{cases} 1, & V_i > \theta \\ 0, & V_i \leqslant \theta \end{cases}. \tag{4}$$

V [mV] is the voltage difference across the membrane potential. W is a gating variable of the channel activation. I_{app} is the applied current stimulus in which the ML model exhibits a periodic and spontaneously firing of the type-I neuron[Fig. 1(a)]. μ is a timescale parameter that can be used to tune the temperature [10]. $C_m = 1.0$, $g_L = 0.5$, $g_K = 2.0$, $g_{Ca} = 1.33$, $V_L = -0.5$, $V_K = -0.7$ and $V_{Ca} = 1.0$. The remaining terms, $m_\infty(V)$, $w_\infty(V)$ and $\tau_w(V)$ are referred to [11].

In the synaptic coupling term, $\bar{i}$ is a counterpart of i. s is a slow variable of the post synaptic conductance. ϵ is synaptic weight where $0 < \epsilon \ll 1$. α is a synaptic relaxation rate. θ is a threshold with $V = 0$[mV]. Once V_i is over the threshold, synaptic conductance s_i influences the partner's membrane potential dynamics.

3 Effects of Temperature on Firing Duration, PRC and Synchrony

The phase reduction method [12] is employed here. Two ML type oscillators with weakly mutual excitation are reduced to equations consisting of the phase degrees

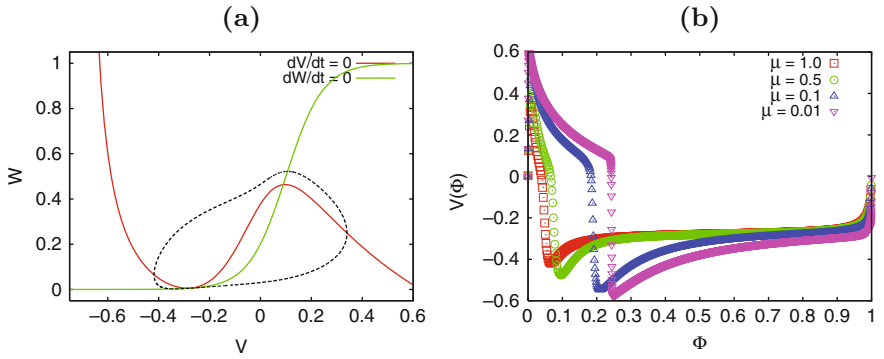


Fig. 1 A phase plane (**a**) and temperature (μ) dependent duration of an action potential (**b**) for the Morris-Lecar model. **a** The trajectory is illustrated counterclockwise as a *dotted line*. *V*- and *W*-nullclines (respectively $[dV/dt] = 0$ and $[dW/dt] = 0$ as described below in Eqs. (1)–(2)) are represented by *red lines*. **b** Time evolution of the membrane potential with different value of μ. Each time evolution t is normalized as ϕ by the corresponding firing frequency measure (Colors will be viewed in online version)

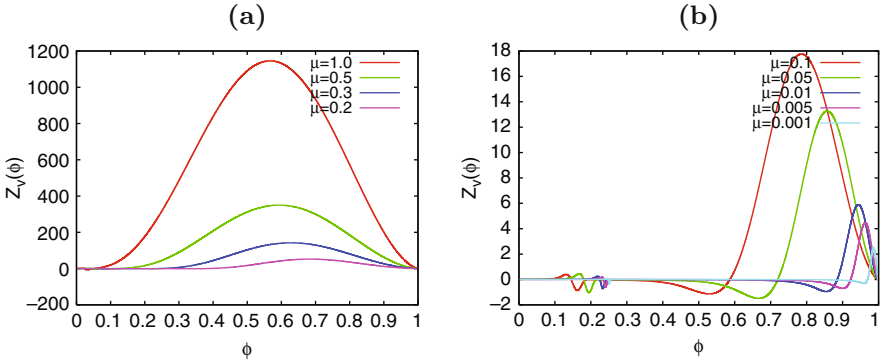


Fig. 2 μ-dependent PRCs for the Morris-Lecar type. **a** Most of PRCs are the type-I. **b** The type-II

of freedom. Simultaneously, the time evolution of the membrane potential is calculated on one cycle (Fig. 1b). Decreasing μ, a periodic firing cycle is much longer with extension of a duration of an action potential. A PRC is also obtained. The so-called type-I PRC, which has almost only positive values (Fig. 2a), is transformed in a change of μ, to reach type-II PRC taking both the positive and negative values (Fig. 2b). Interesting enough, we have shown that the ML model parameterized I_{app} as the type-I neuron takes the type-II RPC, dependent of μ. In linear stability analysis for obtained phase equations, a systematic scheme of synchronization transition is given in terms of μ and α. The synchronization here are defined as ϕ (where $0 \leqslant \phi < 1$) (Fig. 3).

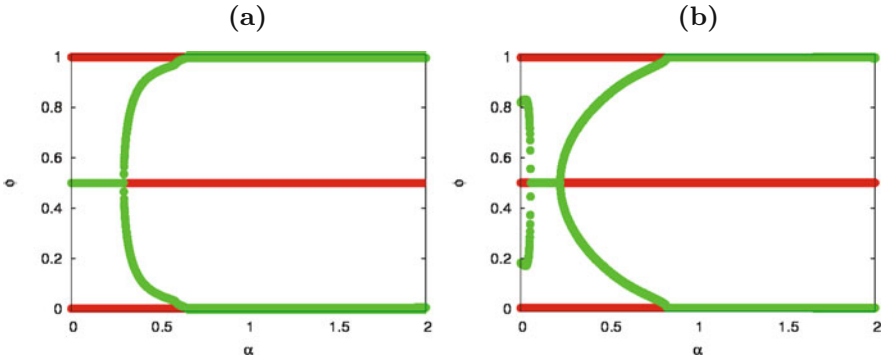


Fig. 3 α-ϕ bifurcations diagram for **a** $\mu = 1.0$ and **b** $\mu = 0.1$. Light *green* and *red lines* respectively represent stable and unstable synchronous states (Colors will be viewed in online version)

4 Conclusion

In a pair system of ML models via excitatory chemical synapses, effects of the temperature μ and the synaptic relaxation rate α on synchrony have been explained. Thermally dependent PRCs were obtained within a theoretical framework of phase reduction analysis. In this analysis, it has also been shown that a duration as well as a firing frequency of the spike are varying in a change of μ. In the past, some researchers suggested that I_{app} or spiking neuron types as well are key factors on synchronization mechanism. However in this work, the dynamical relationship of the temperature on ionic channel mechanisms has been revealed inclusively.

Acknowledgments The author was working in the European Commission-funded project, "Neocortical Daisy Architectures and Graphical Models for Context-Dependent Processing" FP6-2005-015803, in the German Federal Ministry of Education and Research (BMBF) within the "Bernstein Focus: Neurotechnology" through research grant 01GQ0840 and in the Hertie Foundation.

References

1. Braun, H.A., et al.: Int. J. Bifurcat. Chaos. Appl. Sci. Eng. **8** (1998) 881.
2. French, A.S.: J. Comp. Physiol. A. **156** (1985) 817.
3. Beilby, M.J., Coster, H.G.J.: Aus. J. Plant. Physiol. **3**(3) (1976) 275.
4. Cocherová, E.: Proc. ISCR. 132 (2002).
5. Prior, D.J., Grega, D.S.J.: Exp. Biol. **98** (1982) 415.
6. Netoff, T.I., Banks, M.I., Dorval, A.D., Acker, C.D., Haas, J.S., Kopell, N., White, J.A.: J. Neurophysiol. **93** (2005) 1197.
7. Hansel, D., Mato, G., Meunier, C.: Neural. Comput. **7** (1995) 307.
8. Morris, C., Lecar, H.: Biophys. J. **35** (1981) 193.
9. Wilson, H.R.: Spikes, Decisions, and Actions: The Dynamical Foundation of Neuroscience. New York, NY: Oxford University Press (1999).

10. Buia, C.I., Tiesinga, P.H.: J. Neurophysiol. **99** (2008) 2158.
11. Sato, Y.D.: Ph.D. Thesis, Tokyo Institute of Technology (2005).
12. Hoppensteadt, F.C., Izhikevich, E.M.: Weakly Connected Neural Networks. New York, NY: Springer (1997).

Synchronization Transitions of Small-World Neuronal Networks

Wen-Ji Cao and Qing-Yun Wang

Abstract We examine synchronization transitions and patter formation in dependence on the information transmission delay on small-world networks of Morris-Lecar excitable neurons. For the gap junctional coupling, we show that short delays can induce zigzag fronts of excitations, whereas long delays can further detriment synchrony in the network due to a dynamic clustering anti-phase synchronization transition. On the other hand, within the large delay region, which is characterized by anti-phase synchronization and clustering, differences in the network topology do not notably affect the synchrony of neuronal activity.

Keywords Synchronization · Neuronal networks

1 Introduction

Synchronization is an important paradigm in numerous situations that constitute everyday life, and it is only natural that it is in the focus of attention throughout natural as well as social sciences. Since neurons are known to be linked through complex networks [1], extending these findings to specific models of neuronal dynamics [2] seems justified and currently of substantial interest.

At present, we aim to extend the subject by studying pattern formation and synchronization transitions on small-world Morris-Lecar neuronal networks with gap junction, that is characterized with information transmission delay and noise. We report several non-trivial effects induced by finite delay lengths and small-world topology, such as the emergence of zigzag fronts and clustering anti-phase synchronization transition.

W.-J. Cao (✉)
Inner Mongolia Finance and Economics College, Network Information Center, Huhhot 010071, China

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_42,

2 Model Description

The spatiotemporal evolution of the studied neuronal networks, along with additive Gaussian noise and information transmission delay, is governed by the following system [3]:

$$\begin{aligned} \frac{\mathrm{dV_i}}{\mathrm{dt}} &= \frac{1}{\mathrm{C_m}}(I_i^{\mathrm{app}} - I_i^{\mathrm{ion}} - I_i^{syn}) + w\xi_i(t), \\ \frac{\mathrm{dW_i}}{\mathrm{dt}} &= \phi\Lambda(V_i)\left[W_\infty(V_i) - W_i\right] \end{aligned} \tag{1}$$

where $i = 1, \ldots, N$ index the neurons, and V_i and W_i represent the membrane potential and the fraction of open potassium channels, respectively. ξ_i is the Gaussian white noise in the ith neuron with $<\xi_i> = 0$, and $<\xi_i(t)\xi_j(t^{'})> = \delta_{i,j}\delta(t - t^{'})$. w is the noise intensity. I_i^{syn} is the synaptic current, by which, small-world neuronal networks are formed. The values of all parameters and variable that appear in the above equations are given in [3]. We choose parameters such that the Morris-Lecar neuron exhibits the subcritical Hopf bifurcation. Underlying interaction networks for

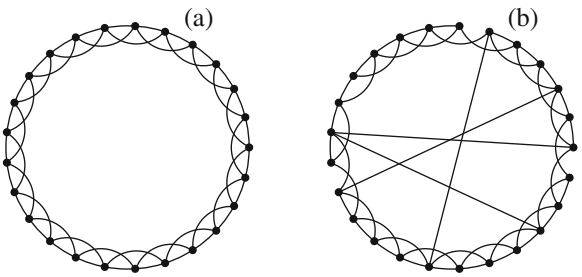


Fig. 1 Examples of considered network topologies. For clarity regarding k and p only 25 vertices are displayed in each panel. **a** Regular ring characterized by $p = 0$ with periodic boundary conditions. Each vertex is connected to its $k = 4$ nearest neighbors. **b** Realization of small-world topology via random rewiring of a certain fraction p of links (in this case 4 out of all 100 were rewired, hence $p = 0.04$)

the ML neurons are obtained by starting from a regular ring with periodic boundary conditions comprising of $N = 300$ vertices, each having $k = 4$ nearest neighbors as shown in Fig. 1a. The parameter p determines the probability of rewiring a link and can occupy any value from the unit interval, whereby $p = 0$ constitutes a regular graph while $p = 1$ results in a random network. For $0 < p < 1$, as exemplified in Fig. 1b.

3 Impact of Delay on Synchronization

The coupling term with delays can be described as follows:

$$I_i^{syn} = g^{syn} \sum_j \varepsilon_{i,j}\left[V_i(t) - V_j(t - \tau)\right]. \tag{2}$$

The coupling strength in this case is set to $g^{syn} = 0.5$ throughout this work, whereby $\varepsilon_{i,j} = 1$ if neuron i is coupled to neuron j and $\varepsilon_{i,j} = 0$ otherwise. τ is the information transmission delay, which will be main parameter on which we focus within this work.

In what follows, the effects of information transmission delay τ and rewiring probability p on patterns formation and synchronization in the examined neuronal networks with gap junction are presented. First, results presented in Fig. 2 illustrate the spatiotemporal dynamics of neurons evoked by different τ on a typical realization of small-world topology given by $p = 0.1$. Initially, in the absence of information transmission delay, neurons can synchronize their spiking as illustrated in Fig. 2a. By short delays, zigzag fronts of excitations can appear as shown in Fig. 2b, which leads to the loss of spiking synchronization. In Fig. 2c, however, it is illustrated that further increases of the delay τ can induce alternative layer patterns, at which excitatory spikes appear alternatively among nearby clusters in space as the temporal dynamics evolves. Hence, this phenomenon can be termed appropriately as a clustering anti-phase synchronization transition induced by an appropriate information transmission delay. Moreover, for larger τ, clustering anti-phase synchronization can also be observed [see Fig. 2d], and it looks like more consistent clustering anti-phase synchronization.

To quantitatively study the degree of spatial synchronization and thus support above visual assessments, we introduce, by means of the standard deviation, a synchronization parameter σ as used earlier in [4], which can be calculated effectively according to

$$\sigma = \frac{1}{T}\int_0^T \sigma(t)dt,\ \sigma(t) = \frac{1}{N}\sum_{j=1}^{N}[V^j(t)]^2 - \left[\frac{1}{N}\sum_{j=1}^{N} V^j(t)\right]^2 \tag{3}$$

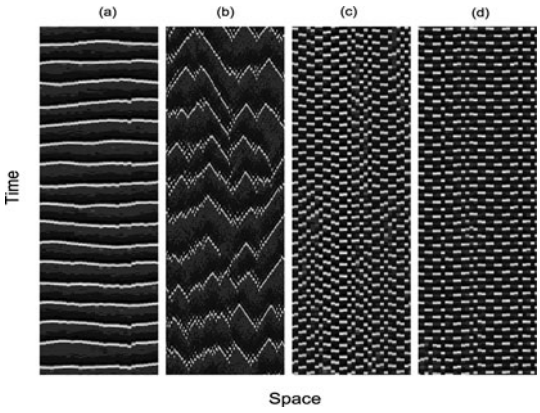


Fig. 2 Space-time plots obtained for different information transmission delays. From *left* to *right* τ equals 0, 4.8, 10 and 12, respectively. Other parameter values are $p = 0.1$ and $w = 2.5$

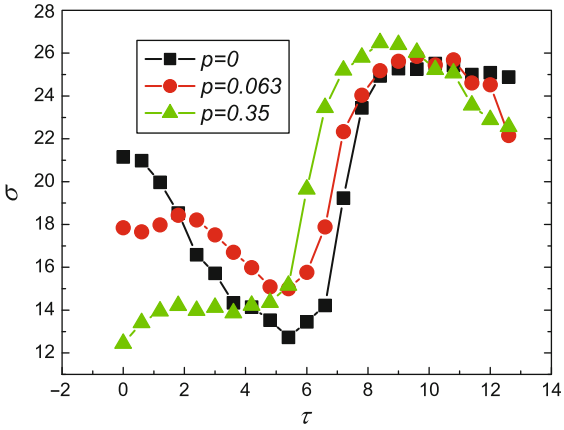


Fig. 3 Dependence of the synchronization parameter σ on τ by different p. Where applicable, other parameters are the same as in Fig. 3

As shown in Fig. 3, we plot σ versus τ for different p. It can be seen that, as the delay increases, σ initially decreases. However, this can not imply the enhancement of spiking synchronization. On the contrary, spiking synchronization is destroyed from the above observations. In fact, it can be seen that for short delays, zig-zag fronts can appear, which destroys synchronization. Decreasing of spiking frequency results in smaller synchronization parameter δ. With delays being increased further, clustering anti-phase synchronization can be found since σ becomes large. We have conducted detailed investigations and found zigzag fronts for the small delays and anti-phase synchronization of nearby clusters for the large values of τ.

4 Conclusion

Character of gap junction is explored in terms of pattern formation and synchronization of the neuronal network as the information transmission delay is varied. It is found that, if the delay is short, neurons within the network can exhibit transitions from zigzag fronts to clustering anti-phase synchronization. Moreover, we show that spatial zigzag fronts can transit to increasingly regular in-phase synchrony as p closes in on the limit constituting random networks. For large delays, however, clustering anti-phase synchronization is independent from the network topology.

Acknowledgments This work was supported by the National Science Foundation of China (Fund Nos. 10702023,10972001 and 10832006).

References

1. Hilgetag, C.C., Burns, G.A., O'Neill, M.A., Scanell, J.W., Young, M.P.: Computational analysis of functional connectivity between areas of primate cerebral cortex. Philos. Trans. R. Soc. B **355** (2000) 91.
2. Hodgkin, A.L., Huxley, A.F.: Quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. **117** (1952) 500.

3. Morris, C., Lecar H.: Voltage oscillations in the barnacle giant muscle fiber. Biophys. J. **35** (1981) 193.
4. Gao, Z., Hu, B., Hu, G.: Stochastic resonance of small-world networks. Phys. Rev. E **65** (2001) 016209.

Bursting Regularity in Complex Networks of Neurons with Chemical Synapses

Xia Shi and Qishao Lu

Abstract In this, we study the bursting dynamics of the complex neural networks with chemical synapses. A quantitative characteristics, width factor, is introduced to describe the bursting dynamics of the neuron. Then we study the effects of the coupling schemes and the network topology on the rhythmic dynamics of the networks. It is concluded that the coupling scheme plays the key role in shaping the bursting types of the neurons. As for the network topology, more links can only change the bursting type of the long bursting neurons, and the short bursting neurons are robust to the link numbers.

Keywords Bursting · Complex networks · Chemical synapses · Neuron

1 Introduction

After the famous papers by Watts and Strogatz [1] and by Barabáasi and Albert [2], much attention has been attracted to the dynamical processes in complex networks in recent years. In biology, neural networks have always been an important subject of research. It is well known that a single neuron in the vertebrate cortex connects to more than 10,000 postsynaptic neurons via synapses forming a complex network [3]. The emergent property of this dynamical system is that a set of neurons will synchronize and fire impulses simultaneously.

Ordering spatiotemporal chaos of small world neural networks with electrical synaptic coupling have been studied extensively [4–7] in recent years. In these studies, they gave some quantitative characteristics to evaluate the spatiotemporal order degree of the neuronal networks. However, they didn't consider the change of the rhythmic dynamics. In this paper, we mainly concentrate on how the properties of the synaptic coupling and the intrinsic properties of the network unit would affect the rhythmic dynamics of the system.

X. Shi (✉)
School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: berthashi@yahoo.com.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_43,

An outline of this paper is given as follows. Some characteristics to evaluate the rhythm dynamics of an isolated bursting neuron are introduced in Section 2. In Section 3, a small world neural network is presented and the effects of the synaptic coupling and the intrinsic properties of the neurons on the rhythm dynamics of the network is studied. Finally, a conclusion is given in Section 4.

2 Rhythm Dynamics of the Bursting Neurons

Bursting neurons present multiple time scales, that is, fast variables relating to the action-potential firing (spiking) and slow processes modulating the fast ones to generate bursting. The discrete time system proposed by Rulkov [8] is

$$\begin{aligned} x_{n+1} &= F_\alpha(x_n, y_n), \\ y_{n+1} &= y_n - \mu(x_n + 1) + \mu\sigma, \end{aligned} \tag{1}$$

where x represents the transmembrane voltage of the neuron and y the slow gating precess. The difference of the time scales between the two subsystems is determined by a sufficiently small value of the parameter μ with $0 < \mu \ll 1$. The nonlinear function $F_\alpha(x, y)$ is responsible for the generation and reset of spikes, and is given by

$$F_\alpha(x, y) = \begin{cases} \frac{\alpha}{1-x} + y, & \text{if } x < 0, \\ \alpha + y, & \text{if } 0 \le x < \alpha + y, \\ -1, & \text{if } x \ge \alpha + y. \end{cases} \tag{2}$$

Depending on the values of the parameters α and σ, the map system reproduces a variety of neural spiking-bursting activities. In this study, we fix $\alpha = 5.0$, $\mu = 0.001$ and control the value of σ to generate various firing patterns.

A burst of spikes is defined as a series of two or more consecutive spikes that have interspike intervals less than a given threshold T_{thresh}. Then the average ratio between the duration of a burst T_b and the subsequent period of refractory time T_r, which is called as width factor and denoted by $f = T_b/T_r$, is introduced to describe the rhythm dynamics of the neuron. If the width factor $f = T_b/T_r > 1$, we call this type of bursts long bursts. Oppositely, the bursts with the width factor $f = T_b/T_r < 1$ are called short ones.

3 Rhythm Dynamics of the Complex Neural Networks

The system of map-based neural network is described as follows

$$\begin{aligned} x_{i,n+1} &= F_\alpha(x_{i,n}, y_{i,n} + \beta_{i,n}), \\ y_{i,n+1} &= y_{i,n} - \mu(x_{i,n} + 1) + \mu\sigma_i, \end{aligned}$$

where $i = 1, 2, \cdots, M$ and $n = 1, 2, \cdots, N$ represent the neuron index and the discrete time series, respectively. The coupling term is modeled by

$$\beta_{i,n} = -g_c \sum_{j=1}^{M} a_{ij} H(x_{j,n} - \theta)(x_{i,n} - \nu), \quad i,j = 1,2, i \neq j \tag{3}$$

through reciprocal excitatory or inhibitory chemical synapses. g_c is the coupling strength and $H(\cdot)$ is the Heaviside step function. The parameter θ is the presynaptic threshold for chemical synaptic interaction and we choose $\theta = 0$ throughout this paper. The constant ν denotes the reversal potential associated with the synapses, which is defined by the nature of the postsynaptic ionic channels. The synapse will be excitatory if ν is higher and inhibitory if ν is lower than the range of $x_{i,n}$. We choose $\nu = 2$ in (3) for excitatory coupling and $\nu = -2$ for inhibitory one. The matrix (a_{ij}) specifies the connection topology: $a_{ij} = a_{ji} = 1$ when a connection exists between neuron i and j, otherwise $a_{ij} = a_{ji} = 0$ and $a_{ii} = 0$ for all i. The small world network is constructed as follows: it starts with a one-dimensional regular ring which comprises $M = 100$ nonidentical neurons and each neuron is connected to its two nearest neighbors. Then the links are randomly added between non-nearest vertices with probability p.

The structural properties of complex networks vary together with the vertices randomly adding probability p. The intrinsic properties of the individual neurons and properties of the synaptic coupling are also important factors in shaping the rhythm dynamics of the network. In the following we choose the parameter value of $\sigma_i \in [0, 0.1]$ for short bursting neurons and $\sigma_i \in [0.16, 0.26]$ for long bursting neurons.

We calculate the average width factor of the network with the change of the coupling strength and p being fixed for two coupling schemes to find the effect of the synaptic coupling on the rhythmic dynamics of the networks. The results are shown in Fig. 1. It is obvious that the evolutions of the width factor are same for different values of adding probability. The neurons in the network will generate long bursts with excitatory coupling and short bursts with inhibitory one, which means that coupling scheme plays the key role in shaping the bursting types of the neurons.

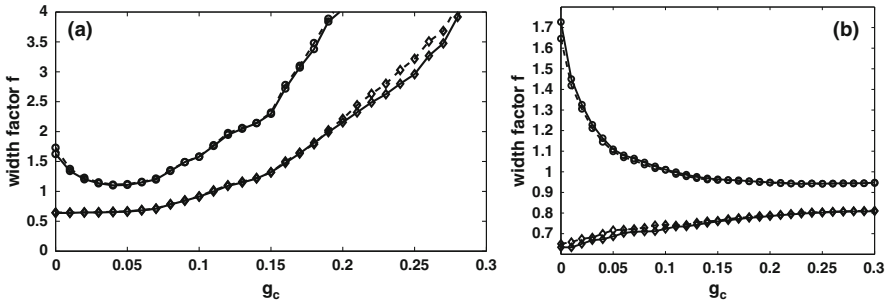


Fig. 1 The evolution of the width factor of the network with the change of the coupling strength g_c. **a** The excitatory coupling and **b** the inhibitory coupling. *Solid* and *dashed* line represent the situations when $p = 0.05$ and $p = 0.1$. ($\circ$) and ($\diamond$) represent the long bursting and short bursting neuronal networks

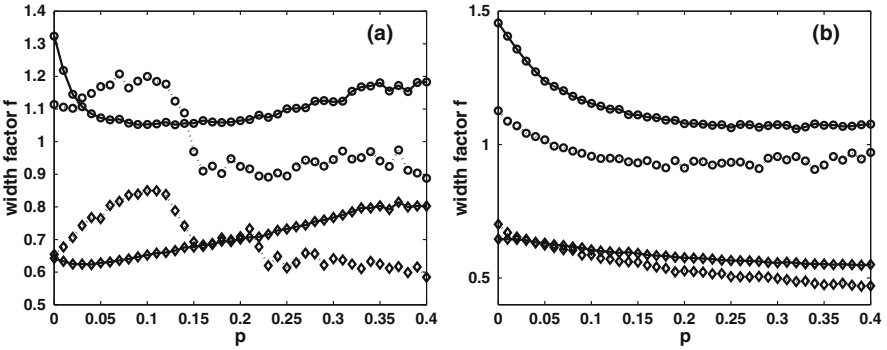


Fig. 2 The evolution of the width factor of the network with the change of the adding probability p. **a** The excitatory coupling and **b** the inhibitory coupling. *Solid* and *dashed* line represent the situations when $g_c = 0.01$ and $g_c = 0.05$. (∘) and (⋄) represent the long bursting and short bursting neuronal networks

Furthermore, we fix the value of g_c and change the value of p to evaluate the effect of the network topology on the rhythmic dynamics of the networks. It is clear from Fig. 2 that for lower coupling strength, adding links cannot modify the bursting types of the network neurons, while adding links transform the long bursting neurons to short ones for higher coupling whatever the coupling scheme is excitatory or inhibitory. In addition, it is concluded that the short bursting type is robust to the network topology.

4 Conclusion

In summary, we have studied the rhythm dynamics of the complex neuronal networks with chemical synapses. We introduced a quantity, width factor, to characterize the bursting dynamics of a single neuron. According to the experimental result, we constructed the small world neuronal network to study the rhythm dynamics of the networks. It is concluded that the coupling scheme plays the key role in shaping the bursting types of the network neurons. As for the network topology, more links can only change the bursting type of the long bursting neurons, and the short bursting neurons are robust to the link numbers.

Acknowledgments This work was supported by the National Natural Science Foundation of China (No.10802012 and No. 10872014).

References

1. Watts, D.J. Strogatz, S.H.: Collective dynamics of "small-world" networks. Nature **393** (1998) 440–442.
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286** (1999) 509–512.

3. Gerstner, W., Kistler, W.M., eds.: Spiking Neuron Models. Cambridge: Cambridge University Press (2002).
4. Wang, M., Hou, Z., Xin, H.: Ordering spatiotemporal chaos in small-world neuron networks. Chemphyschem **7** (2006) 579–582.
5. Gong, Y., Xu, B., et al.: Ordering spatiotemporal chaos in complex thermosensitive neuron networks. Phys. Rev. E. **73** (2006) 046137.
6. Wei, D.Q., Luo, X.S.: Ordering spatiotemporal chaos in discrete neural networks with small-world connections. Europhys. Lett. **78** (2007) 68004.
7. Shen, Y., Hou, Z., Xin, H.: Transition to burst synchronization in coupled neuron networks. Phys. Rev. E. **77** (2008) 031920.
8. Rulkov, N.F.: Modeling of spiking-bursting neural behavior using two-dimensional map. Phys. Rev. E. **65** (2002) 041922.

Burst Synchronization in a Small-World Neuronal Network with Coupling Delays

Jiaoyan Wang, Qishao Lu, Fang Han, and Zhuoqin Yang

Abstract In this paper, burst synchronization is studied in a heterogeneous small-world neuronal network of non-identical Hindmarsh-Rose (HR) neurons with coupling delays. We study the effect of time delays and coupling strength. It is observed that the network with optimal delay can achieve both burst synchronization and phase synchronization much easily when the coupling strength is sufficiently large. On the contrary, the network cannot achieve burst synchronization or phase synchronization even with sufficiently large coupling strength and some delays. It is also found that the network can achieve burst synchronization but not phase synchronization with some delays and coupling strength. Moreover, we study the effect of noise on burst synchronization.

Keywords Burst sycnronization · Small-world neuronal network · Coupling delays

1 Introduction

The observation of synchronous neural activities in the central nervous system suggests that neural activity is a cooperative process of neurons and synchronization plays a vital role in mechanisms of information processing and information preface within different brain area [1]. Controlling chaos and synchronization has been well studied during the last two decades, and a variety of approaches have been presented to implant this task, ranging from dissipative systems to conservative systems.

Although the real structure of the neuronal network of human brain is not clear yet, small-world properties, such as dense clustering and short average path length, have been found in some neuronal networks. Thus small-world structure may be instructive for the essence of neuronal networks of human brain. Considering the

J. Wang (✉)
Department of Dynamics and Control, Beihang University, Beijing 100191, China

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_44,

fact that the signal will take some time to pass through a finite distance to arrive at its destination due to limited speed and then induce a time-delay in receiving signals. In neural networks, the transmission delay is a sum of axonal, synaptic, and dendritic delays. Moreover, it is reported that axons can generate time delays as large as 300 ms [2].

In this paper, we study the influence of network topology and parameters on burst synchronization in a small-world neuronal network of nonidentical chaotic HR neurons with delays. This article is organized as follows. Section 2 introduces a model of small-world HR neuronal network. Section 3 presents the main results of burst synchronization and spatiotemporal patterns in detail.

2 Model

The single HR neuron model proposed by [3] can be written in the form:

A neuronal network composed of electrically coupled HR neurons is given by

$$\begin{aligned}
\frac{dx_i}{dt} &= y_i - ax_i^3 + bx_i^2 - z_i + I_i - g\sum_{j=1}^{N}\frac{a_{ij}}{\sum_{k=1}^{N} a_{ik}} f(x_i, x_j^{\tau}),\\
\frac{dy_i}{dt} &= c - dx_i^2 - y_i,\\
\frac{dz_i}{dt} &= r[s(x_i - \chi) - z_i],\\
f(x_i, x_j^{\tau}) &= x_i(t) - x_j(t-\tau)
\end{aligned} \tag{1}$$

$i = 1, 2 \ldots \mathrm{N}$, $a = 1$, $b = 3$, $c = 1$, $d = 5$, $s = 4$, $r = 0.006$ and $\chi = -1.6$, where N is the overall number of neurons. We set the external current of the i-th neuron $I_i = 3.1 + 0.001^{*}i$. The coupling matrix A is determined by the network structure and the element a_{ij} takes the value 1 or 0, depending on whether or not there exists electrical coupling between neurons. Here, we use the Newman–Watts small world structure in system (1). We start with a regular ring of $N = 20$ neurons, each coupled diffusively to its two nearest neighbors. Then we add shortcuts between pairs of nodes with probability p.

3 Main Results and Discussion

3.1 Phase Synchronization and Burst Synchronization

Phase synchronization of neurons is actually spike synchronization, which means all neurons emit spikes within fixed phase differences while the amplitudes stay uncorrelated. The phase of the i-th neuron can be defined by the Hilbert transformation of the membrane potential x_i. The Hilbert transformation of signal $x(t)$ can

be expressed as $x_h(t) = (1/\pi) \int_{-\infty}^{+\infty} (x(\tau)/(t-\tau))d\tau$, and the instantaneous phase of signal $x(t)$ is defined as $\phi(t) = \arctan\left(\frac{x_h(t)}{x(t)}\right)$.

According to this definition, whenever a neuron fires a spike, the phase of the neuron increases 2π. Then the average phase difference over time for all neurons is used to indicate the phase synchronization of network, that is, $\Delta\phi(t) = \frac{1}{N-1}\sum_{i=2}^{N} |\phi_i(t) - \phi_1(t)|$. When $\max|\Delta\phi(t)| < 2\pi$, phase synchronization of the small-world neuronal network is thought to happen.

Similarly, the burst phase of the i-th neuron can be defined as [4]

$$\phi_i(t) = 2\pi k + \frac{2\pi(t-t_k)}{t_{k+1}-tk}, \quad t_k < t < t_{k+1}, \tag{2}$$

where t_k is the time at which the k-th burst starts. Let

$$\Delta\phi(t) = \tfrac{1}{N-1}\textstyle\sum_{i=2}^{N} |\phi_i(t) - \phi_1(t)|.$$

When $\max|\Delta\phi(t)| < 2\pi$, burst synchronization of the small-world neuronal network is thought to happen.

It is observed that the network with optimal delay e.g. $\tau = 30$ can achieve phase synchronization (see Fig. 1a) and burst synchronization (see Fig. 1b) easily when the coupling strength is sufficiently large. On the contrary, the network cannot achieve phase synchronization (see Fig. 1c) or burst synchronization (see Fig. 1d) with some sufficiently large coupling strength with some delays, e.g. $\tau = 51$. It seems that the network can achieve burst synchronization more easily with some smaller coupling strength than with some large coupling strength when $\tau = 51$. It is found that the network can achieve burst synchronization (see Fig. 1b, d) but cannot achieve phase synchronization (see Fig. 1a, c) when $\tau = 1$ and $\tau = 0$. Moreover, burst synchronization is enhanced when $\tau = 1$ compared with $\tau = 0$. So the network can achieve global burst synchronization much easier than phase synchronization. Furthermore, we can see that time delays can enhance or destroy synchronization from Figs. 1 and 2.

3.2 *The Effect of Noise on Phase Synchronization and Burst Synchronization*

To make the model more realistic, we add Gaussian white noise ξ_i to the i-th neuron with $< \xi_i >= 0$ and $< \xi_i\xi_j(t') >= D\delta_{i,j}\delta(t-t')$, where D represents the noise intensity. Arbitrary small noise can destroy the exact synchrony of neurons [5]. From Figs. 3 and 4, phase synchronization and burst synchronization survives at small noise. It is clear that burst synchronization survives at bigger noise than phase synchronization.

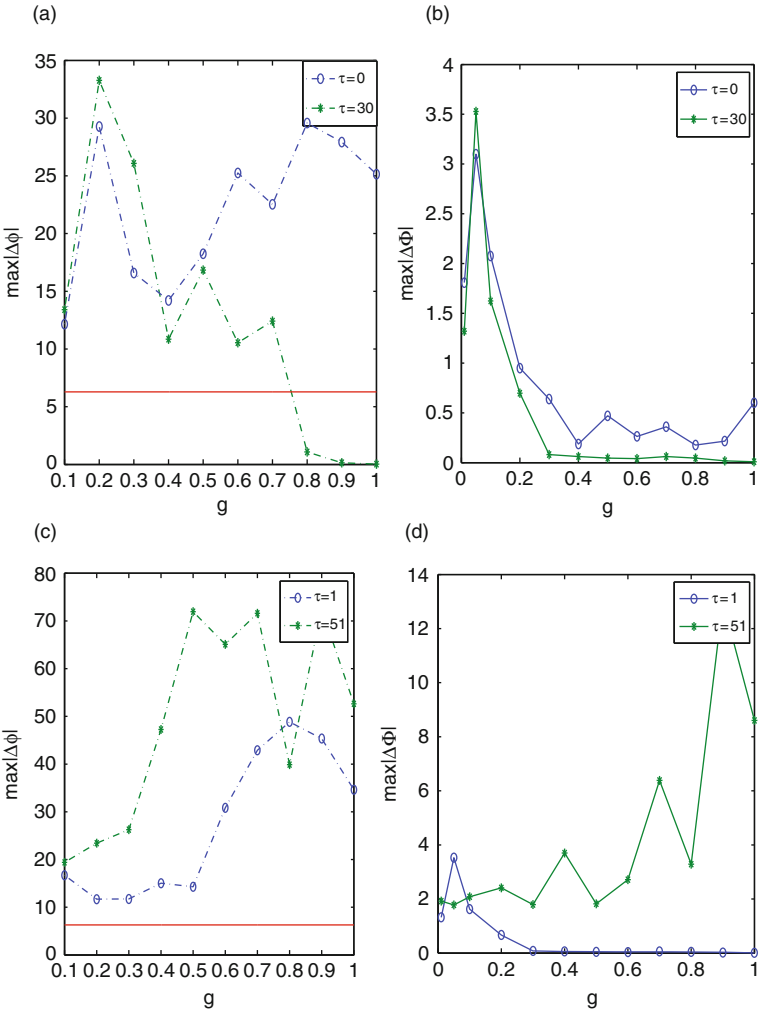


Fig. 1 **a** and **c** Phase synchronization; **b** and **d** Burst synchronization with different coupling strength at $p = 0.08$, $N = 20$

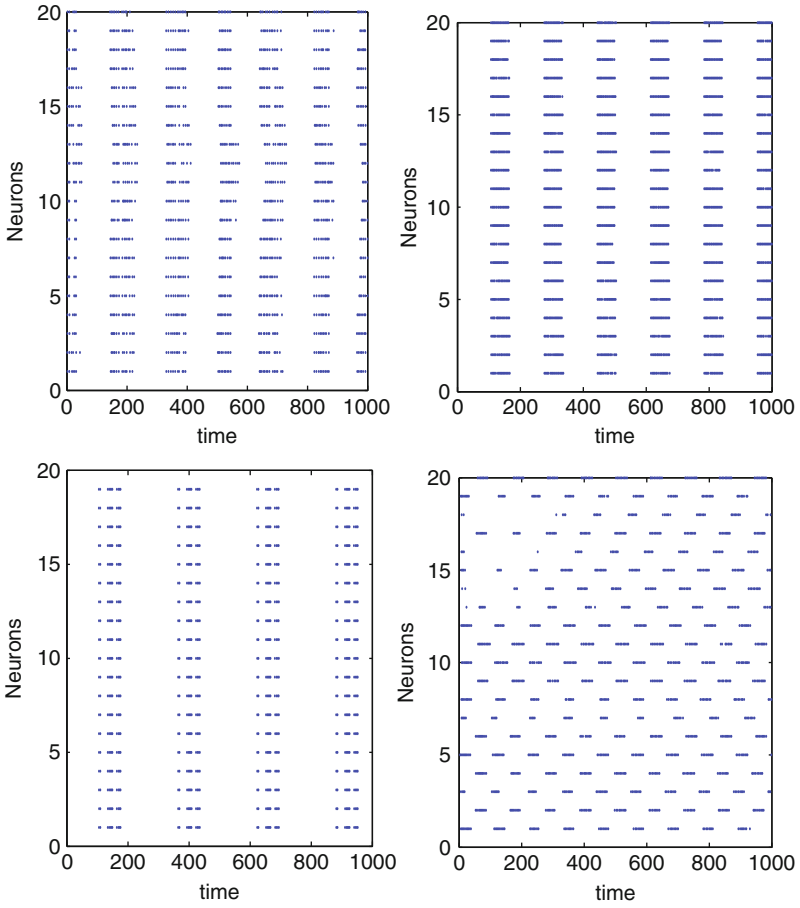


Fig. 2 Spatial patterns of the network for $\tau = 0,\ 1,\ 30,\ 51$ at $p = 0.08$, $g = 1$, $N = 20$

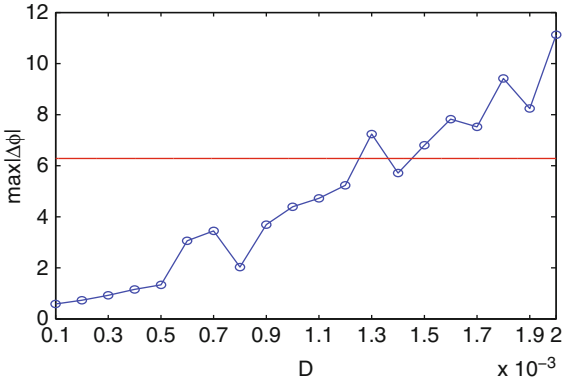


Fig. 3 Phase synchronization for different noise strengths at $\tau = 39$, $p = 0.08$, $g = 0.7$, $N = 20$

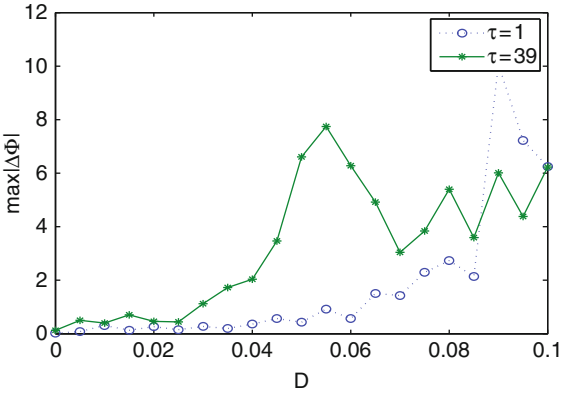


Fig. 4 Burst synchronization for different noise strengths at $p = 0.08,\ g = 0.7,\ N = 20$

Acknowledgements This work was supported by the National Natural Science Foundation of China (No.10702002 and 10872014).

References

1. Singer, W.: Synchronization of cortical activity and its putative role in information processing and learning. Annu. Rev. Physiol. **5** (1993) 349.
2. Aboitiz, F., et al.: Fiber Composition of the human corpus callosum. Brain Res. **598** (1992) 143.
3. Hindmarsh, J.L., Rose, R.M.: A model of neuronal bursting using three coupled first order differential equations. Proc. R. Soc. Lond. Ser. B. **221** (1984) 87.
4. Ivanchenko, M.V., Osipov, G.V., Shalfeev, V.D., Krths, J.: Phase synchronization in ensembles of bursting oscillators. Phys. Rev. Lett. **93** (2004) 134101.
5. Dhamala, M., Jirsa, V.K., Ding, M.: Enhancement of neural synchrony by time delay. Phys. Rev. Lett. **92** (2004) 074104.

Correlation-Induced Phase Synchronization in Bursting Neurons

Xiufeng Lang and Qishao Lu

Abstract We investigate phase synchronization of uncoupled bursting neurons, caused by spatially correlated noise. It is found that the degree of synchronization between non-identical neurons decreases monotonously when increasing partially correlated noise intensity, but undergoes a minimum with the common noise intensity increasing.

Keywords Neuron · Phase synchronization · Noise · Bursting

1 Introduction

In the past decades, bursting behavior has been found in many neurobiological systems [1]. Recently, bursting synchronization and its role in the neurobiological system has attracted more and more attention. Especially, noise-induced synchronization of chaotic systems was demonstrated [2]. In spite of these studies, it remains unclear that how noise influences the synchronization in uncoupled neurons with the dynamics of chaotic bursting. In addition, some experiments have observed that noise could have significant effects on the dynamics of neural systems [3]. Herein, we intensively study the effects of spatially correlated noise on phase synchronization (PS) in two non-identical bursting neurons. It is found that the degree of PS monotonously decreases with the partially correlated noise, while the degree non-monotonously varies with the common noise. The variety might be attributed to the interplay between noise correlation and parameter mismatch.

X. Lang (✉)
Department of Dynamics and Control, Beihang University, Beijing 100191, China

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_45,

2 Model

The dynamics of the i-th Hindmarsh-Rose neuron [4] ($i = 1, 2$) is described by the following differential equations:

$$\frac{dx_i}{dt} = y_i - ax_i^3 + bx_i^2 - z_i + I_i + D\xi_i \tag{1}$$

$$\frac{dy_i}{dt} = c - dx_i^2 - y_i \tag{2}$$

$$\frac{dz_i}{dt} = r[s(x_i - x_0) - z_i] \tag{3}$$

where x_i is the membrane potential, y_i is associated with the fast current, and z_i with the slow current. Here $a = 1.0,\ b = 3.0,\ c = 1.0,\ d = 5.0,\ s = 4.0,\ r = 0.006$, $x_0 = -1.6$, and I_i is the external current input. The spatially correlated noise with the intensity D consists of two components, that is, $\xi_i(t) = \sqrt{R}e(t) + \sqrt{1-R}\eta_i(t)$, where $e(t)$, a common Guassian noise, and $\eta_i(t)$ is a local Guassian noise. The control parameter R measures the noise correlation between the neurons.

3 Results and Discussion

To investigate the effect of correlated noise on PS, a phase variable for spike trains of the variable x is defined as follows [5]:

$$\varphi(t) = 2\pi\left(k + \frac{(t - \tau_k)}{(\tau_{k+1} - \tau_k)}\right) \tag{4}$$

where $\tau_k < t < \tau_{k+1}$, and τ_k the k-th firing time of neuron. The most common approach to study phase synchronization behavior in stochastic systems is to compute the distribution of cyclic phase difference $P(\Delta_c\varphi)$ on $[-\pi, \pi]$, where $\Delta_c\varphi = \Delta\varphi \bmod (2\pi)$.

To quantify the degree of PS, we introduce the entropy of distribution of cyclic phase difference as follows [6]:

$$H = -\sum_{i=1}^{M} P_i \ln P_i \tag{5}$$

where M is the number of bins and P_i is the probability when the cyclic phase difference $\Delta_c\varphi$ is in the i-th bin. Normalizing H in [0, 1] by the maximal entropy $H_m = \ln M$ of the uniform distribution, the synchronization index is calculated by:

$$\rho_{ps} = \frac{H_m - H}{H_m} \tag{6}$$

Here $\rho_{ps} = 0$ corresponds to a uniform distribution (no synchronization) and $\rho_{ps} = 1$ corresponds to a Dirac-like distribution (perfect synchronization).

In the absence of noise, we introduce a mismatch into the external currents of two neurons $I_1 = 3.1$ and $I_2 = 3.12$, where the two neurons both exhibit chaotic bursting. Figure 1a, c show the distribution of cycle phase difference $P(\Delta_c\varphi)$ for $D = 0$, 1 and 10 with $R = 1$ (i.e., the common noise), respectively. In Fig. 1a, high peaks of $P(\Delta_c\varphi)$ (e.g., $P(\Delta_c\varphi) > 0.05$) appear in many bins of $\Delta_c\varphi$ in the deterministic limit ($D = 0$), which indicates that phase locking between the bursting neurons occurs at multiple phase differences, characterizing imperfect phase synchronization in multiple-time-scaled system discussed in [6]. As the noise intensity D increases to 1 (Fig. 1b), high peaks of $P(\Delta_c\varphi)$ appears only at less bins of $\Delta_c\varphi$ and height of the highest peak is close to that in which demonstrates that compared with those in Fig. 1a, less phase differences could be locked between the neurons, and moreover, the strength of stochastic phase locking hardly changes at the phase difference for the highest peak. When D increases to 10 (Fig. 1c), high peaks occurs in the least bins of $\Delta_c\varphi$, but the highest peak are much higher than those in Fig. 1a, b. This illustrates that compared with the case of weak noise, strong noise makes the neurons stochastically phase locking at less phase differences, but the strength of phase locking is improved significantly at the phase difference for the highest peak.

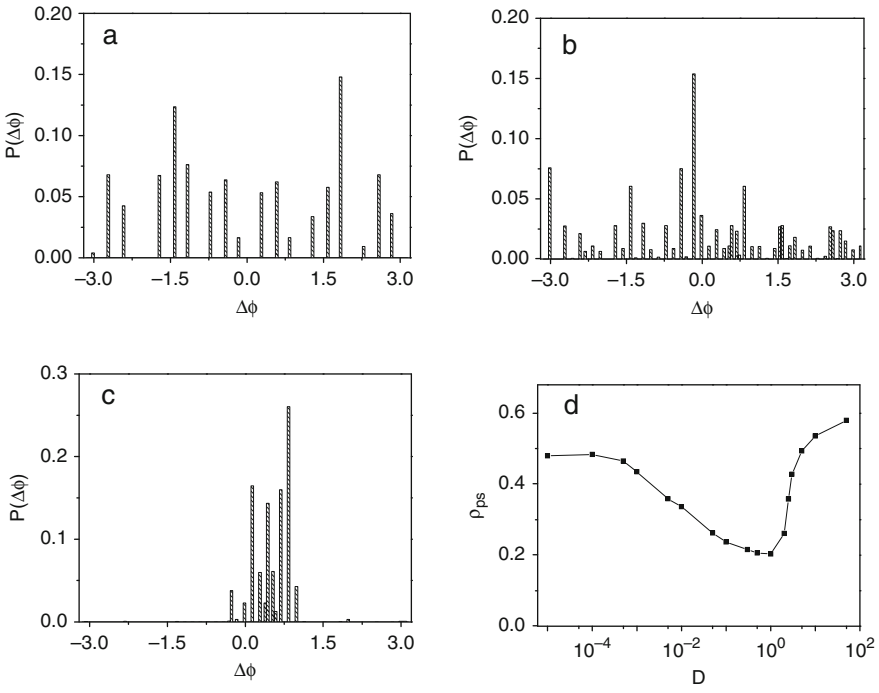


Fig. 1 Distributions of cyclic phase difference for (**a**) $D = 0$, (**b**) $D = 1$ and (**c**) $D = 10$ with $R = 1$; (**d**) phase synchronization index ρ_{ps} as a function of D for $R = 1$

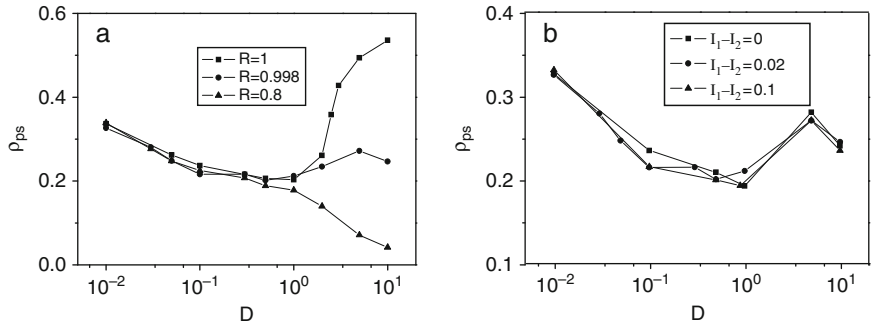


Fig. 2 **a** Phase synchronization index ρ_{ps} as a function of D for $R = 0.8$, 0.998 and 1. **b** ρ_{ps} versus D for $I_1 - I_2 = 0,\ 0.02$ and 1

Aforementioned results imply that when the common noise intensity increases, the degree of PS might first decrease at low noise level and then increase at higher noise level, which is further demonstrated by curves of ρ_{ps} versus D in Fig. 1d. It is clearly seen from the figure that ρ_{ps} undergoes a minimum with D increasing, indicating that the degree of PS between the two neurons become the weakest at certain common noise intensity.

However, the situation for the partially correlated noise ($0 < R < 1$) is quite different. As shown in Fig. 2a, ρ_{ps} increases slightly for $R = 0.998$, and descends greatly with noise intensity for $R = 0.8$, which indicates that the degree of PS monotonously decreases with the noise intensity increasing. The various effects of partially correlated noise and common noise on PS might be attributed to the interplay between noise correlation and parameter mismatch. Figure 2b plots ρ_{ps} versus D for three values of difference between the external currents of two neurons. It is found that for certain noise intensity and correlation, the ρ_{ps} hardly changes with the parameter heterogeneity, demonstrating that parameter mismatch has little effect on the degree of PS between two non-identical neurons. It should be noted that these phenomena also holds for other values of R (the data are not shown here). Therefore, it is possible that the noise correlation changes the mean external inputs, especially for the strong noise, which leads to the enhancement of PS for the common noise and decrease of PS for the partially correlated noise.

4 Summary

In this paper we investigated the nontrivial effects of spatially correlated noise, consisting of a common Guassian noise and a local one, on phase synchronization of two unconnected and non-identical bursting neurons. It is found that the degree of synchronization is minimized at some common noise intensity, while is decreased monotonously with the partially correlated noise intensity increasing. It is further

demonstrated that for strong noise, noise correlation could change the degree of synchronization more greatly than parameter mismatch.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 10872014).

References

1. Krahe, R., Gabbiani, F.: Nat. Rev. **5** (2004) 13.
2. R. Toral, C.R., Mirasso, E., Hernández-Garcia, O.: Piro Chaos. **11** (2001) 665.
3. Doiron, B. Lindner, B., Longtin, A., Maler, L., Bastian, J.: Phys. Rev. Lett. **93** (2004) 048101.
4. Hindmarsh, J.L., Rose, R.M.: Proc. R. Soc. Lond. Ser. B. **221** (1984) 87.
5. Pikovsky, A., Rosenblum, M., Kurths, J.: Synchronization-A Unified Approach to Nonlinear Science. Cambridge, UK: Cambridge University Press (2001).
6. Yuan, Z., Zhang, J., Zhou, T.:Phys. Rev. E. **78** (2008) 031901.

Synchronization of Small-World Neuronal Networks with Synapse Plasticity

Fang Han, Qishao Lu, Xiangying Meng, and Jiaoyan Wang

Abstract Synchronization of small-world neuronal network with synapse plasticity is explored in this paper. The variation properties of synapse weights are studied first, and then the effects of synapse learning coefficient, the coupling strength and the adding probability on synchronization of the neuronal network are studied respectively. It is shown that appropriate learning coefficient is helpful for improving synchronization, and complete synchronization can be obtained by increasing the coupling strength and the adding probability.

Keywords Neuronal network · Small-world · Synchronization · Synapse plasticity

1 Introduction

Synchronization is thought to play a key role in intercommunications among neurons. With the developing of the theories of complex networks, people pay more attention on the synchronization of complex neuronal networks, especially small-world ones. Most of the research focuses on non-weighted or constant-weighted neuronal networks, however, the weights of synapses among neurons keep changing in the growth of neurons and in the studying and memorizing processes. In other words, synapse can learn, which is also called synapse plasticity. In this paper, the synchronization of small-world neuronal networks with synapse plasticity is explored numerically.

F. Han (✉)
School of Information Science and Technology, Donghua University, Shanghai 201620, China; Department of Dynamics and Control, Beihang University, Beijing 100191, China
e-mail: yadiahan@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_46,

2 The Mathematical Model

By using the Hindmarsh-Rose neuronal model [1] and the Newman-Watts small-world strategy [2], we can set up electrically coupled neuronal network. The number of neurons is set as $N = 100$.

As for the synapse plasticity, we define the variation equation for the weight w_{ij} between neuron i and neuron j as following [3]:

$$\Delta w_{ij} = L * \arctan[x_i(x_j - x_i w_{ij})] \tag{1}$$

where L is a positive synapse learning coefficient and x_i is the membrane potential of neuron i.

Then the neuronal network with synapse plasticity can be expressed as:

$$\begin{aligned} \dot{x}_i &= y_i - a x_i^3 + b x_i^2 - z_i + I + \sigma \sum_{j=1,\, j\neq i}^{N} g_{ij} w_{ij}(x_j - x_i), \\ \dot{y}_i &= c - d x_i^2 - y_i, \\ \dot{z}_i &= r[s(x_i + \chi) - z_i], \\ \Delta w_{ij} &= L \arctan[x_i(x_j - x_i w_{ij})], \qquad i = 1, 2, \cdots, N. \end{aligned} \tag{2}$$

where σ is the coupling strength, $\mathbf{G} = \{g_{ij}\}_{N\times N}$ is the connectivity matrix of the network, I is the external current. The parameters in the system are set as $a = 1$, $b = 3, c = 1, d = 5, s = 4, r = 0.006, \chi = 1.6, I = 3$. As ions in channels are transmitted mutually, bi-directional non-equal connections are adopted here. That is, if there is a connection between neuron i and neuron j ($i \neq j$), then there exists a weight $w_{ij} \neq w_{ji}$, if not, then $w_{ij} = w_{ji} = 0$. And let $w_{ii} = 0$, which means that there is no self-connection for all the neurons in the network.

3 Simulation Results

The variation properties of synapse weights under the influence of the synapse learning coefficient is studied first. Figure 1 shows the variations of membrane potentials, connection weights and connection strengths between two arbitrarily chosen connected neurons in the network. The connection strength of neuron i is defined as $s_i(t) = \sum_{j=1}^{N} w_{ij}(t)$.

It can be seen from Fig. 1 that, if both of the two neurons are in rest, the connection weights and the connection strengths between them do not change; if neuron i is excited and neuron j is in rest, the connection weight from neuron i to neuron j is strengthened, and the connection weight from neuron j to neuron i is weakened. That means, when ion currents conduct from one neuron to another by the connection channels between them, the positive connection weight would be strengthened and the negative connection weight would be weakened. These phenomena conform with actual variations of synapses during the excitement transmission process among real neurons.

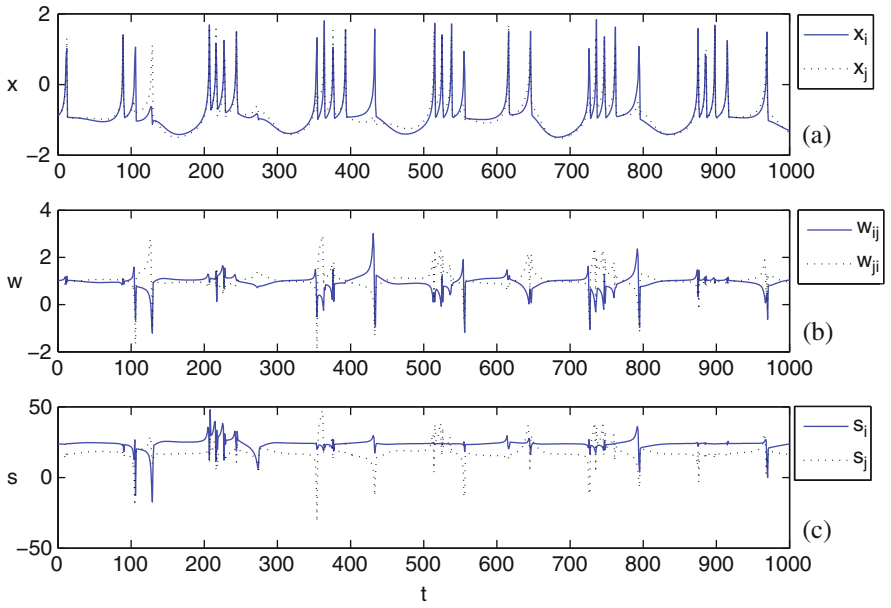


Fig. 1 The variations of membrane potentials, connection weights and connection strengths between two arbitrarily chosen connected neurons ($L = 8, \sigma = 0.05, p = 0.1$): **a** x_i and x_j are the membrane potentials of neuron i and neuron j respectively; **b** w_{ij} and w_{ji} are the positive and negative weights between neuron i and neuron j respectively; **c** s_i and s_j are the connection strengths of neuron i and neuron j respectively

Then we discuss synchronization properties of the network under the influence of synapse plasticity. We use average synchronization error as characteristic measure, which can be represented as

$$e_{\text{mean}} = \text{mean}(\langle e_i \rangle), \langle e_i \rangle = \langle \sqrt{(x_i - x_1)^2 + (y_i - y_1)^2 + (z_i - z_1)^2} \rangle$$

where $\langle \cdot \rangle$ represents average on time, $i = 2, 3, \cdots, N$.

By numerical simulation, the relationship of the average synchronization error and the synapse learning coefficient can be obtained as in Fig. 2.

It can be seen that, the average synchronization error decreases sharply first, then oscillates around 0.5 finally, which means that the increase of synapse learning coefficient is helpful for synchronization, but can not make the network achieve synchronization. It implies that in neuronal networks, the learning function of electric synapses must be neither too strong nor too weak, but be appropriate to maintain the synchronizability.

The variations of the average synchronization error with the increasing coupling strength and the adding probability are shown in Fig. 3a, b respectively.

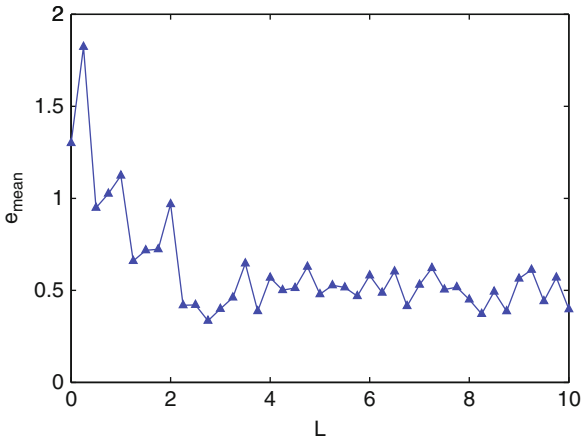


Fig. 2 The relationship of the average synchronization error e_{mean} and the learning coefficient $L(\sigma = 0.05, p = 0.1)$

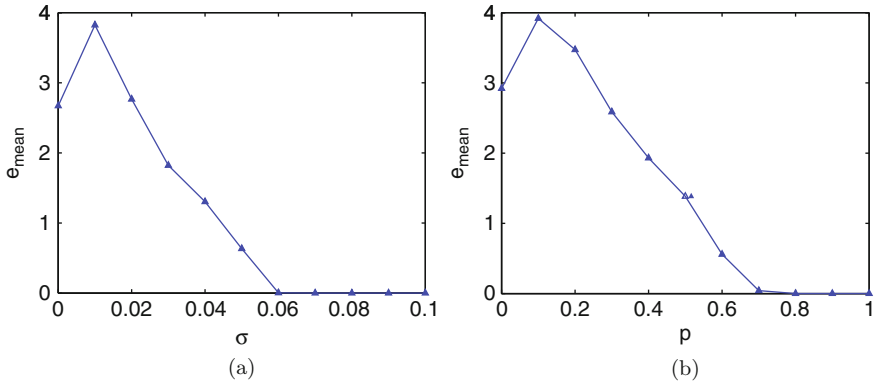


Fig. 3 **a** The relationship of the average synchronization error e_{mean} and the coupling strength $\sigma(L = 2, p = 0.1)$; **b** The relationship of the average synchronization error e_{mean} and the adding probability $p(L = 2, \sigma = 0.01)$

It can be seen from Fig. 3a that, when the coupling strength increases to $\sigma = 0.06$, the average synchronization error decreases to zero, which means the network achieves complete synchronization. Hence, increasing the coupling strengths among neurons can finally make the network synchronize completely. It also can be seen from Fig. 3b that, when the adding probability increases to $p \approx 0.7$, the average synchronization error decreases to zero. So, in neuronal networks with synapse plasticity, the introduction of shortcuts also can improve synchronizability of neuronal networks.

4 Conclusion

Based on that actual biological neuronal networks have the properties of small-world connectivity and the connection strengths among neurons change dynamically, we study the synchronization of electrically coupled neuronal networks with synapse plasticity. The variation properties of connection weights are studied first. Then the effects of the synapse learning coefficient, the coupling strength and the adding probability on synchronization are studied respectively. It is shown that appropriate learning coefficient is helpful for improving the synchronizability of the network, and increasing the coupling strength and the adding probability can finally make the neuronal network synchronize completely.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No.10872014).

References

1. Hindmarsh, J.L., Rose, R.M.: A model of neuronal bursting using three coupled first order differential equations. Proc. R. Soc. Lond. Ser. B. **221** (1984) 87–102.
2. Newman, M.E.J., Watts, D.J.: Scaling and percolation in the small-world network model. Phys. Rev. E. **60** (1999) 7332–7342.
3. Zheng, H.Y., Luo, X.S., Lei, Wu: Excitement and optimality properties of small-world biological neural networks with updated weights (in Chinese). Acta. Phisica. Sin. **57**(6) (2008) 3380–3384.

A Modeling Study of Glutamate Release of NMDA Receptors

Justin Blackwell, Jianzhong Su, and Ege Kavalali

Abstract In this chapter we will use a mathematical model to simulate spontaneous and evoked neurotransmission to determine how receptor location on presynaptic density can activate distinct sets of NMDA receptors and signal independently to the postsynaptic side. We will also examine how the release rate of the neurotransmitter Glutamate will affect post-synaptic currents by varying the fusion pore size at the presynaptic side.

Keywords Synapse modeling · Neurotransmission · Independent signaling · NMDA receptors · Diffusion equation · Chemical synapses

1 Introduction

Spontaneous synaptic vesicle fusion is a salient feature of all synapses [1, 2]. These random release events typically activate receptors within a single postsynaptic site and give rise to miniature postsynaptic currents, and therefore, they have been extremely instrumental in analysis of unitary properties of neurotransmission. Spontaneous release typically occurs with a rate of 1–2 vesicles per minute per release site [3–5], whereas evoked release at individual synapses can occur at an extremely high rate (>100 vesicles per second) [6].

To determine how neurons distinguish evoked and spontaneous neurotransmission we examine the possibility that spontaneous and evoked neurotransmission activate non-overlapping NMDA receptor populations.

J. Blackwell (✉)
Department of Mathematics, The University of Texas at Arlington, Arlington, TX 76019, USA

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_47,

2 Model

Our glutamate diffusion model followed the approach previously used by Nielsen et al. [7] and first simulated isotropic diffusion of 4,000 glutamate molecules [8] released from a point source. We then simulated release of the glutamate molecules through a vesicle by addition of two compartments [9] that one modeled the vesicle (40 nm by 40 nm) and the other represented the fusion pore (width of 2 and 10 nm's).

The standard thermodiffusion equation is applicable for the glutamate concentration $C_{i,j,k} = C\left(x_i, y_j, z_k\right)$, in both the synaptic cleft and synapse, as follows:

$$\frac{\partial C}{\partial t} = D_{\text{glut}} \left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C^2}{\partial y^2} + \frac{\partial^2 C}{\partial z^2} \right) \quad . \tag{1}$$

The diffusion constant D_{glut} takes different values depending on the location. When simulating the vesicular diffusion the value of D_{glut} is 0.15 μm^2/ms and 0.0375 μm^2/ms within the fusion pore. Once the molecules leave the vesicle the value of D_{glut} takes a value of 0.4 μm^2/ms inside the cleft and 0.75 μm^2/ms outside the cleft.

To determine the opening probability, P_{open}, of an individual receptor a model consisting of three closed and two open states (3C2O) is used [10]. The glutamate concentrations across the cleft are included by adding two coupled states, C^M and C^U.

$$\text{(M)} \qquad C^U \underset{58}{\overset{39[G]}{\rightleftharpoons}} C^M \underset{116}{\overset{19[G]}{\rightleftharpoons}} C_1 \underset{173}{\overset{150}{\rightleftharpoons}} C_2 \underset{2{,}412}{\overset{902}{\rightleftharpoons}} C_3 \underset{1{,}283}{\overset{4{,}467}{\rightleftharpoons}} O_1 \underset{526}{\overset{4{,}630}{\rightleftharpoons}} O_2 \tag{2}$$

$$\text{(L)} \qquad C^U \underset{60}{\overset{34[G]}{\rightleftharpoons}} C^M \underset{120}{\overset{17[G]}{\rightleftharpoons}} C_1 \underset{161}{\overset{127}{\rightleftharpoons}} C_2 \underset{2{,}610}{\overset{580}{\rightleftharpoons}} C_3 \underset{2{,}167}{\overset{2{,}508}{\rightleftharpoons}} O_1 \underset{662}{\overset{3{,}449}{\rightleftharpoons}} O_2 \tag{3}$$

On the post synaptic cleft there are 16 receptors, arranged in a 4 by 4 array using row-major ordering. The concentration time-course and P_{open} are calculated at each receptor. Two receptors are used as release sites, R6 for central release and R16 for release at the boundary.

3 Results

From the results we conclude that peak open value is not as sensitive to the pore sizes as is distance from release site. Thus, factors such as varying the diffusion constant and the distance from receptor to release site is more relevant in determining spontaneous versus evoked transmission. By varying the release rates, there is a noticeable difference in max concentrations of glutamate, as shown in Fig. 1. Instant release assumes instantaneous release of 4,000 molecules from a point source. We then compared release from a 10 nm and 2 nm fusion pore.

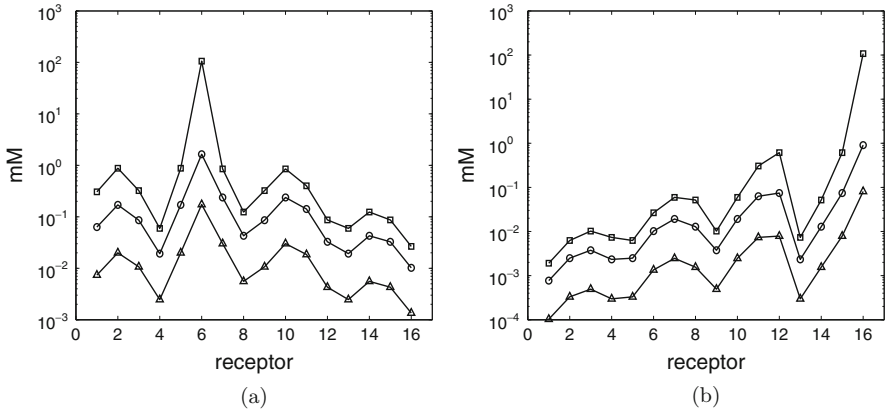


Fig. 1 The charts depict maximum concentrations of neurotransmitters (glutamate) at receptor locations, where the *square* represents a scenario of instantaneous release at cleft, the *circle* is for releasing through a 10 nm fusion pore, and the *triangle* is for releasing through a 2 nm fusion pore. **a** Release of neurotransmitter near center of synapse at receptor 6. **b** Release of neurotransmitter near edge of synapse at receptor 16

When estimating the open probabilities of all receptors in response to glutamate release onto R6, the site of evoked release as shown in Fig. 2, if all receptors are in the same mode, then R6 has a 4.5 fold increase to that of R16. However, this difference decreases to a 1.9 fold difference if R16 is in M-mode and R6 is in L-mode. This difference is compatible with experimental results in which blocking evoked NMDA responses results in a small reduction of spontaneous NMDA activity [11].

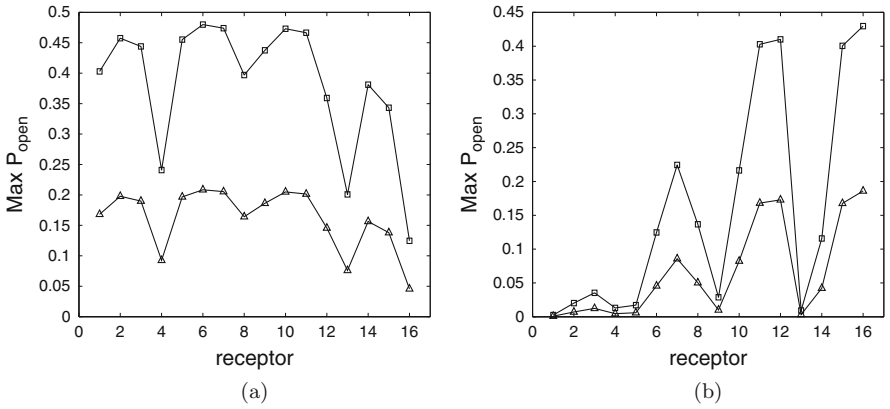


Fig. 2 The charts depict maximum channel opening probability at receptor locations. **a** Central instantaneous release. **b** Boundary instantaneous release. *Square* is M-mode and *triangle* is L-mode

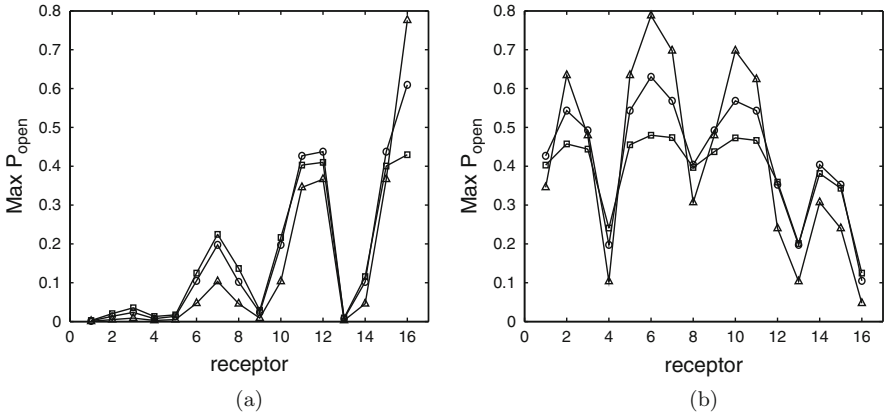


Fig. 3 The charts represent maximum channel opening probability by peak values for M-mode kinetics at receptor locations. **a** Release of neurotransmitter near center of synapse. **b** Release of neurotransmitter near edge of synapse. The *square* represents instantaneous release, the *circle* is for 10 nm, and the *triangle* is for 2 nm vesicular release

This illustrates that a large synapse of 0.36 μm^2 (with R6 near the center for evoked neurotransmission, and R16 for spontaneous neurotransmission) can accommodate two non-overlapping domains that give rise to independent signaling.

Figure 3 compares the max values of P_{open} with various release rates. The peak value increases as the size of the fusion pore decreases due to prolonged exposure to the neurotransmitter

References

1. Fatt, P., Katz, B.: Spontaneous subthreshold activity at motor nerve endings. J. Physiol. **117** (1952) 109–128.
2. Del Castillo, J., Katz, B.: Quantal components of the end-plate potential. J. Physiol. **124** (1954) 560–573.
3. Geppert, M., Goda, Y., Hammer, R.E., Li, C., Rosahl, T.W., Stevens, C.F, Südof, T.C.: Synaptotagmin I: A major Ca2+ sensor for neurotransmitter at a central synapse. Cell. **79** (1994) 717–727.
4. Murthy, V.N., Stevens, C.F.: Reversal of synaptic vesicle docking at central synapses. Nat. Neurosci. **2** (1990) 503–507.
5. Sara, Y., Virmani, T., Deák F, Liu, X., Kavalali, E.T.: An isolated pool of vesicles recycles at rest and drives spontaneous neurotransmission. Neuron **45** (2005) 563–573.
6. Saviane, C., Silver, R.A.: Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse. Nature **439** (2006) 983–987.
7. Nielsen, T.A., DiGregario, D.A., Silver, R.A.: Modulation of glutamate mobility reveals the mechanism underlying slow-rising AMPAR EPSCs and the diffusion coefficients in the synaptic cleft. Neuron **24** (2004) 757–771.
8. Xu-Friedman, M.A., Regehr, W.G.: Ultrastructural contributions to desensitizational at cerebellar mossy fiber to granule cell synapse. J. Neurosci. **23** (2003) 2182–2192.

9. Choi, S., Klingauf, J., Tsien, R.W.: Fusion pore modulation as a presynaptic mechanism contributing to expression of long-term potentiation. Phil. Trans. R. Soc. Lond. **358** (2003) 695–705.
10. Popescu, G., Robert, A., Howe, J.R., Auerbach, A: Reaction mechanism determines NMDA receptor response to repetitive stimulation. Nature **430** (2004) 790–793.
11. Atasoy, D., Ertunc, M., Moulder, K.L, Blackwell, J., Chung, C., Su, J., Kavalali, E.T.: Spontaneous and evoked glutamate release activates two populations of NMDA receptors with limited overlap. J. Neurosci. **28** (2008) 10151–10166.

Experimental Discovery of a Complex Bifurcation Scenario from Period 1 Bursting to Period 1 Spiking

Minghao Yang, Qiaohua Zheng, Chunling Wei, Huaguang Gu, and Wei Ren

Abstract A complex bifurcation scenario from period 1 bursting to period 1 spiking containing period increment bifurcation scenario in bursting patterns, simulated in theoretical model in previous studies, awaited experimental demonstration. In this Letter, the bifurcation scenario was discovered in an experimental neural pacemaker. The characteristics of the bifurcation scenario, the period increment bifurcation scenario and the transition from bursting to spiking pattern were also identified. The results demonstrate the reality of the theoretical prediction of firing pattern transition regularity.

Keywords Firing pattern · Bifurcation scenarios · Period increment bifurcation · Bursting · Spiking

1 Introduction

In recent years, nonlinear dynamics in neural firing rhythm have attracted more and more attention. The dynamics of neural firing patterns and firing pattern transitions are of fundamental importance to elucidate neural coding mechanism. In the last two decades, various firing patterns including bursting and spiking pattern, and various firing pattern bifurcation scenario from bursting to spiking pattern were simulated in a series of neuronal models, such as Chay model and Hindmarsh-Rose (HR) model [1–6]. For example, many different kinds of bifurcation scenarios from period 1 bursting to period 1 spiking via simple procedure and complex procedure were simulated.

M. Yang (✉)
College of Life Science, Shaanxi Normal University, Xi'an, Shaanxi 710062, China
e-mail: minghaoyang@yahoo.com.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_48,

Guided by the previous theoretical studies simulated in Chay model and HR model, we discovered various bifurcation scenarios of firing patterns in the experimental neural pacemaker [4–6]. Most of the bifurcation scenarios from period 1 bursting to period 1 spiking were discovered in the experiment [6]. Unfortunately, only a part of the bifurcation scenario from period 1 bursting to period 1 spiking containing period increment bifurcation scenario in bursting patterns, such as from period 1 bursting to period 7 bursting [5] were observed in the experimental neural pacemaker, but the complete procedure of the bifurcation scenario from period 1 bursting to period 1 spiking, suggested to be the universality in biological rhythm [2], have been not discovered in the experiments for a long time [5, 6]. In this Letter, we report the recent discovery of this bifurcation scenario in the experimental neural pacemaker.

2 Theoretical Model and Simulation Results

Chay model [1, 2] contains the H–H type ionic channel dynamics, verified to be closely relevant to our experimental neural pacemaker, shown as follows:

$$\frac{dV}{dt} = g_I m_\infty^3 h_\infty (v_I - V) + g_{kv}(v_k - V)n^4 + g_{kc}\frac{C}{1+C}(v_k - V) + g_I(v_I - V) \quad (1)$$

$$\frac{dn}{dt} = \frac{n_\infty - n}{\lambda n(\alpha_n + \beta_n)} \quad (2)$$

$$\frac{dC}{dt} = \rho\left[m_\infty^3 h_\infty (v_C - V) - k_C C\right] \quad (3)$$

where t is the time. V is the membrane potential. n is the activation probability of potassium channel. C is the intracellular concentration of calcium ion. The explicit interpretation of Chay model was given in previous study [1]. In this Letter, the parameter values are $v_\mathrm{I} = 100$ mV, $v_k = -75$ mV, $v_1 = -40$ mV, $g_1 = 7$ mS/cm^2, $1,800$ mS/cm^2, $g_\mathrm{kc} = 10$ mS/cm^2, $g_\mathrm{kv} = 1,700$ mS/cm^2, $\rho = 0.27$ and $\lambda_\mathrm{n} = 235$. v_c is the reversal potential of calcium ion, chosen as bifurcation parameter.

A Gaussian white noise, $\xi(t)$, reflecting fluctuation of environments, is directly added to the right hand of Eq. (1) to form the stochastic Chay model. $\xi(t)$ possesses the statistical properties as $\langle\xi(t)\rangle = 0$ and $\langle\xi(t)\xi(t')\rangle = 2D\delta(t - t')$, where D is the noise density and $\delta(\cdot)$ is the Dirac δ-function. The Chay model is solved by mannella numerical integrate method with integration time step being 10^{-3} s. Upstrokes of the voltage reached the amplitude of –25.0 mV were counted as spikes.

To be compared to the experimental results, the bifurcation scenario from period 1 bursting to period 1 spiking is also simulated. The deterministic and stochastic dynamics of the bifurcation scenario were particularly shown in previous studies [5] and introduced simply as follows. When $\lambda_\mathrm{n} = 235$ and v_c is decreased, a bifurcation scenario from period 1 bursting to period 1 spiking containing period increment bifurcation scenario in bursting pattern is simulated. When v_c is larger than 101, the firing pattern is bursting. When v_c is less than 101, the firing pattern is spiking. The

"shrink" phenomenon induced by the transition from bursting pattern to spiking can be seen from the bifurcation scenario when v_c is about 101. When v_c is decreased from 485 to 101, a period increment bifurcation scenario is exhibited in the bursting pattern. The corresponding bifurcation scenario in the stochastic Chay model, exhibiting an overlap near each bifurcation point of bursting pattern. The overlap is the noise induced stochastic bursting. For each bifurcation point, the period k bursting is changed into period $k + 1$ bursting via the stochastic bursting whose behavior is stochastic transition between period k burst and period $k + 1$ burst ($k = 1, 2, 3, 4, 5, 6, 7, 8, 9$).

3 Experimental Model and Results

Experiments were performed on an experimental neural pacemaker, formed at the injured site of adult male SD rat sciatic nerve subjected to chronic ligature [4–6]. After surgical operation and a survival time of 6–12 days, the previously injured site was exposed and perfused with 34°C Kreb's solution. The spontaneous firing of a single individual fiber was recorded with a Powerlab system (Australia) with sampling frequency being 10.0 kHz. The time intervals between the maximal values of the successive spikes were recorded seriatim as interspike interval (ISI) series.

Extracellular calcium concentration ($[Ca^{2+}]_o$) was as bifurcation parameter and was changed from 1.2 mmol/L to zero. The bifurcation scenario from period 1 bursting to period 1 spiking containing period increment bifurcation scenario in bursting pattern was discovered in 7 experimental neural pacemakers when $[Ca^{2+}]_o$ was decreased. In this Letter, one of 7 bifurcation scenarios was studied as a representative. Period increment bifurcation scenario in bursting pattern was period adding bifurcation scenario with stochastic bursting near each bifurcation point, similar to those of the stochastic Chay model. The procedure was from period 1 bursting to stochastic bursting, to period 2 bursting, to stochastic bursting, to period 3 bursting, to stochastic bursting, to period 4 bursting, to stochastic bursting, to period 5 bursting, to firing pattern with multiple spikes in a burst (the number of spikes in a burst was not fixed because noise or the fluctuation of $[Ca^{2+}]_o$), and finally to period 1 spiking. Near each period increment bifurcation point, the behavior of the stochastic bursting lying between period k bursting and period $k + 1$ bursting was stochastic transition between period k burst and period $k + 1$ burst ($k = 1, 2, 3, 4, 5$). The stochastic firing pattern was studied particularly in previous studies. The most important characteristic of this bifurcation scenario was the "shrink" phenomenon corresponding to the transition from bursting to spiking pattern. In the bifurcation scenario, the bursting pattern was changed into spiking pattern via a procedure that the long quiescence between bursts in the bursting pattern disappear, and then the firing pattern became tonic spiking. The disappearance of the long quiescence looks like a "shrink" in ISI series. The frequency of period 1 spiking is much higher than that of the period 1 bursting.

4 Discussion and Conclusion

Based on the results of this Letter and our previous studies [6], various bifurcation scenarios from period 1 bursting to period 1 spiking via simple or complex procedure, were presented in the experiment. The various bifurcation scenarios provide framework to recognize the relationship between different firing patterns.

Acknowledgments This work was supported by project of NSFC under Grant Nos 10772101 and 30770701, as well as the National High Technology Research and Development Program of China under Grant No. 2007AA02Z310.

References

1. Chay, T.R.: Chaos in a three-variable model of an excitable cell. Physica D. **16** (1985) 233–242.
2. Chay, T.R., Fan, Y.S., Lee, Y.S.: Bursting, spiking, chaos, fractals, and universality in biological rhythms. Int. J. Bifurcat. Chaos. **5** (1995) 595–635.
3. Fan, Y.S., Holden, A.V.: Bifurcations, burstings, chaos and crises in the Rose-Hindmarsh model for neuronal activity. Chaos Solit. Fract. **3** (1993) 439–449.
4. Ren, W., Hu, S.J., Zhang, B.J., Wang, F.Z., Gong, Y.F., Xu, J.X.: Period-adding bifurcation with chaos in the interspike intervals generated by an experimental neural pacemaker. Int. J. Bifurcat. Chaos. **7** (1997) 1867–1872.
5. Gu, H.G., Yang, M.H., Li, L., Liu, Z.Q., Ren, W.: Dynamics of autonomous stochastic resonance in neural period adding bifurcation scenarios. Phys. Lett. A. **319** (2003) 89–96.
6. Li, L., Gu, H.G., Yang, M.H., Liu, Z.Q., Ren, W.: A series of bifurcation scenarios in the firing transitions in an experimental neural pacemaker. Int. J. Bifurcat. Chaos. **14** (2004) 1813–1817.

Exploring the Asymmetrical Cortical Activity in Motor Areas Using Support Vector Machines

Guangxin Huang, Huafu Chen, Weishuai Lv, and Qian Dai

Abstract A linear SVM with different inputs, i.e., the average volume within a task block and a single volume of a task block was used to explore the asymmetrical cortical activity in motor areas based on a vivo fMRI data of hand-grip experiment. Our results showed that a SVM with different inputs could get the discriminative activations between the left and right hand grips.

Keywords Functional magnetic resonance imaging · Support vector machine · Pattern recognition methods · Maximum-margin hyperplane

1 Introduction

Support vector machines (SVMs) [1, 2] are a set of multivariate methods. SVM had been applied to fMRI data processing previously [3–5]. SVM can extract a unique brain discrepancy map which is directly related to the measured effect difference between different experimental brain conditions [3–5]. Mourão-Miranda et al. [3, 4] used individual time points of whole brain as input to SVM to obtain the brain discrepancy maps. Wang et al. [5] have shown that the statistical test of discrimination maps may be carried out using random effects analysis (RFX) [6] and permutation testing (PMU) [7]. In the present study, SVM with different inputs was used to explore the asymmetrical cortical activity in motor areas between the left and right hand grip from individual time points of whole brain of single-subject functional magnetic resonance imaging (fMRI) data.

H. Chen (✉)
Key laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, School of Applied Mathematics, University of Electronic Science and Technology of China, Chengdu 610054, China
e-mail: chenhf@uestc.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_49,

2 Methods

2.1 SVM

Support vector machine (SVM) [1, 2] is a data classification algorithm. In the linear case the method consists of maximizing the margin between the support vectors and the hyperplane. Let the training data, a set of points of the form:

$$D = \left\{(x_i, y_i)|x_i \in R^M, y_i \in \{-1, 1\}, i = 1, ..., N\right\},$$

where y_i is either 1 or –1, indicating the classes A and B to which the point x_i belongs. Each x_i is an M-dimensional real vector, corresponding to an individual time point of whole brain scan). The main idea of SVM is to find the maximum-margin hyperplane dividing the points having $y_i = -1$ from those having $y_i = 1$.

In the general case of overlapping classes, the problem of finding the optimal separating hyperplane that maximizes the distance to the nearest training points of the two classes is defined as the following optimization problem:

$$\text{Minimize } \frac{1}{2}\left\|W\right\|^2 + C\sum_i \xi_i \tag{1}$$

$$\text{subject to } y_i(w \cdot x_i - b) \geq 1 - \xi_i, 1 \leq i \leq N \tag{2}$$

$$\text{and } \xi_\text{i} \geq 0, 1 \leq i \leq N. \tag{3}$$

where $\xi_\text{i} \geq 0, 1 \leq i \leq N$ are slack variables that account for training errors and C is a positive real constant appearing only as an additional constraint on the Lagrange multipliers. Eqs. (1–3) can be translated into their unconstrained dual form. By using Lagrangian multiplier method, the solution can be found as [2]:

$$w = \sum_{i=1}^{N} \alpha_i y_i x_i \tag{4}$$

where $\alpha_i \geq 0$ is the Lagrangian multiplier, which is constrained by $\alpha_i y_i = 0$. The points with $\alpha_i > 0$ are called "support vectors", which lie in the supporting hyperplane. For any other points, $\alpha_i = 0$.

The weight vector w describes the importance of each feature in discriminating between the two classes. In the study, w represents a map of the most discriminating regions, we call it as a spatial discrimination map to differ the previous researches [3–5]. Given two classes, task 1 and task 2, with the labels +1 and –1, a positive value in the map w means that this voxel had higher activity during task 1 than task 2 in the training examples that contributed most to the overall classification, i.e., the support vectors; negative value indicates lower activation during task 1 than task 2.

2.2 Permutation Test, Dimensionality Reduction and Pre-processing

As a nonparametric technique, the permutation test [7] has been previously applied to fMRI data analysis [5]. Here, this technology was used to determine the threshold of the spatial discrimination map. Under the null hypothesis of no relationship between the class labels and the global structure of the observations in data space, the class labels were permuted 2,000 times randomly to get all the p-values. All significant voxels with p-values <0.001 were displayed by the values in the weight vector$w$in the spatial discrimination map. PCA was used for the purpose of data compression without losing any information. A detailed description of PCA we refer to [3, 4]. Before the dimensionality reduction, experimental data the experimental data were preprocessed using SPM2. Reconstructed images were corrected for slice timing effects and motion artifacts, as well as transformed to a standard space at $3 \times 3 \times 3$ mm^3. Spatial smoothing with an isotropic Gaussian kernel of 8 mm FWHM was also performed.

3 Applying SVM with Different Inputs to fMRI Data

3.1 Material

The experiment, stimuli and fMRI data were detailed in our early work [8].

3.2 SVM with the Input of the Average Volume

Let x_i be the average volume of the ith task (the left or right hand grip) block, y_i is the task label, i.e., $y_i = 1$ for the right hand grip, $y_i = -1$ for the left hand grip. Let $< x_i, y_i >$ be the input to SVM, then we can get the weight vector w. Figure 1a presents the discrimination maps of a single subject on the left and right hand grips with the average volume as input to a SVM, all voxels with p-value < 0.001 are shown in a color scale corresponding to the values in the weight vector

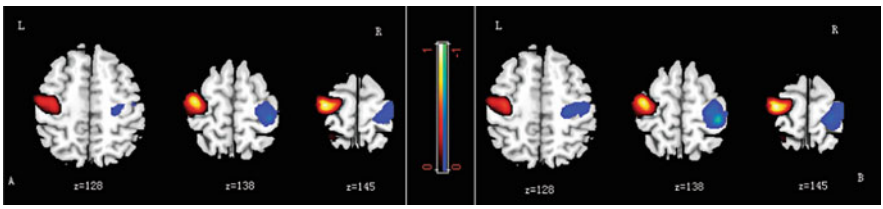


Fig. 1 Spatial discrimination maps of a single subject on the left and right hand grips with the average input (**a**) and a single input (**b**) to SVM. All voxels with p-value < 0.001 are shown in *color scale* (Colors will be viewed in online version)

w. The color scale identifies the most discriminating regions for each time point (light/dark blue for negative values, i.e., relatively more activation for the left hand grip, and red/orange for positive values, i.e., relatively more activation for the right hand grip).

3.3 SVM with the Input of a Single Volume

Let x_i be a single volume within a task (right or left hand grip) block, y_i is the task label, i.e., $y_i = 1$ for right hand grip, $y_i = -1$ for left hand grip. Let $< x_i, y_i >$ be the input to SVM, we can get the weight vector w. Figure 1b shows the discrimination maps of a single subject on left and right hand grips with the average volume as input to a SVM, all voxels with p-value < 0.001 are shown in a color scale corresponding to the values in the weight vector w.

4 Conclusion

This paper mainly applied SVM with different inputs, i.e., the average volume of the left or right hand grip task block and a single volume within the left or right hand grip task block to explore the asymmetrical cortical activity in motor areas. The SPM t-contrast (right hand grip > left hand grip) was also done.

Acknowledgements This work was partly supported by NSFC, 9082006 and 30770590, Key research project of science and technology of MOE (107097) and 863 Program No: 2008AA02Z4080.

References

1. Vapnik, V.: Statistical Learning Theory. New York, NY: JohnWiley and Sons Inc. (1998).
2. Burges, C.-J.C.: A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov. **2** (1998) 121–167.
3. Mourão-Miranda, J., Bokde, A.-L.W., Born, C., Hampel, H., Stetter, S.: Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data. NeuroImage. **28** (2005) 980–995.
4. Mourão-Miranda, J., Reynaud, E., McGlone, F., Calvert, G., Brammer, M.: The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. NeuroImage. **33** (2006) 1055–1065.
5. Wang, Z., Childress, A.R., Wang, J., Detre, J.A.: Support vectormachine learning-based fMRI data group analysis. NeuroImage. **36** (2007) 1139–51.
6. Penny, W., Holmes, A., Friston, K.: Random effects analysis. In Frackowiak, R., Friston, K., Frith, C., Dolan, R., Friston, K., Price, C., Zeki, S., Ashburner, J., Penny, W., eds.: Human Brain Function. 2rd edn. New York, NY: Academic Press (2003).
7. Nichols, T., Holmes, A.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. **15** (2002) 1–25.
8. Zeng, L., Chen, H.F., Ouyang, L., Yao, D.Z., Gao, J.H.: Quantitative analysis of asymmetrical cortical activity in motor areas during sequential finger movemen. Magn. Reson. Imaging. **25** (2007) 1370–1375.

Strange Non-chaotic Attractors in Noisy FitzHugh-Nagumo Neuron Model

Guang-Jun Zhang, Jue Wang, Jian-Xue Xu, Hong Yao, and Xiang-Bo Wang

Abstract Strange Non-chaotic attractors in noisy FHN neuron model periodically driven are researched in this paper. Here we show, based on a nonlinear dynamical analysis and numerical evidence, that under the perturbation of weak noise strange non-chaotic attractor can be induced in FHN neuron model. And the mechanism of strange non-chaotic attractor is related to transitions among chaotic attractor, periodic-3 attractor and chaotic saddle in two sides of crisis point of system respectively.

Keywords Strange non-chaotic attractor · Chaotic saddle · Periodic-3 attractor · Chaotic attractor

1 Introduction

It is well known that neurons work in a noisy environment, and it is therefore of great interest to study how information is encoded and transmitted when neurons work in such a noisy environment [1]. The subject of strange non-chaotic attractors (SNAs) has attracted continuous interest in the nonlinear and statistical physics community [2, 3]. Here "strange" refers to the nontrivial, complicated geometry of the attractor, and "nonchaotic" indicates that the maximum Lyapunov exponent of the attractor is non-positive and there is thus no sensitive dependence on initial conditions [2, 3]. In principle, strange non-chaotic attractors occur in all dissipative dynamical systems that exhibit the period-doubling route to chaos, where the attractors formed at the accumulation points of period-doubling cascades are fractal sets with zero Lyapunov exponent [3]. In some neuron model which leads to

G.-J. Zhang (✉)
School of Life Science and Technology, Xi'an Jiaotong University, Xi'an Shaanxi, PR China;
School of Aerospace, Xi'an Jiaotong University, Xi'an Shaanxi, PR China
e-mail: zhanggj3@126.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_50,

chaos by the route of periodic-doubling bifurcation, such as autonomous, periodically driven and quasiperiodically driven dynamical system, as typical dissipative dynamical systems, there should exist robust strange non-chaotic attractors. But this has not attracted interest of researchers. The results of neurophysiology experiment show that non-periodic firing neurons are more sensitive to stimulis than periodic firing neurons [4]. For some chaotic spike neuron model, the strange non-chaotic attractor is of particular importance in the research of encoding and transmission of neural information in noisy environment. There probably are some relations between some phenomena in neuron spike, such as chaotic synchronization, and strange non-chaotic attractors. So in the paper, strange non-chaotic attractors in periodically driven FHN neuron model under the perturbation of weak noise are researched.

2 The Bifurcation Characteristic of the Periodically Driven FHN Neuron Model

We consider the periodically driven FitzHugh-Nagumo neuron model in the following form Ref. [5]:

$$\begin{cases} \varepsilon \frac{dv}{dt} = v(v-a)(1-v) - w \\ \frac{dw}{dt} = v - dw - b + r\sin\beta t \end{cases} \tag{1}$$

The variable v is the fast voltagelike variable and w is the slow recovery variable. Throughout the paper we fix the values of the constants to $\varepsilon = 0.02$, $d = 0.78$, $a = 0.5$, $r = 0.27$, $\beta = 14.0$. And parameter b is as controlling parameter. A firing or spike is considered to occur if the variable v has a positive-going crossing of the firing threshold vth, chosen as vth $= 0.5$. According to Ref. [6], by modified generalized cell mapping, the nonlinear dynamics global characteristic of FHN neuron model is obtained. The bifurcation figure of system is shown in Fig. 1. As shown in Fig. 1, it can be seen that with the change of b system leads to chaos by the route of periodic-doubling bifurcation. When b is some certain value crisis in FHN neuron model occurs. And when $b = 0.2404$, boundary crisis occurs. A chaotic saddle in FHN neuron model appears after boundary crisis occurs. In a certain parameter region before boundary crisis there coexist a chaotic attractor and a period-3 attractor. But after boundary crisis chaotic attractor disappears, the system displays a period-3 spiking, and a chaotic saddle remains.

3 The Strange Non-chaotic Attractor in FHN Neuron

In this paper, the Gaussian distributed white noise $\xi(t)$ is used for perturbing the system. The mean and autocorrelation function are as follows respectively:

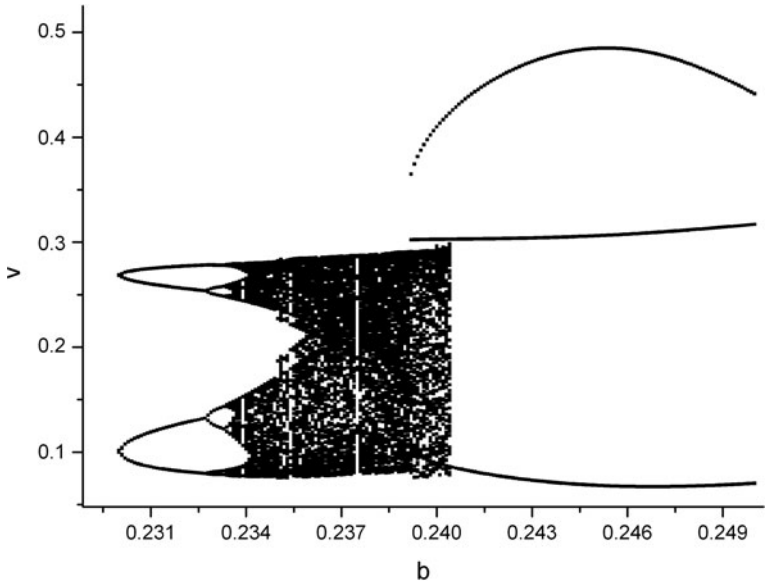


Fig. 1 The bifurcation characteristic of FHN neuron model

$$\begin{cases} < \xi(t) >= 0 \\ < \xi(t)\xi(s) >= 2D\delta(t-s) \end{cases} \tag{2}$$

where D is the noise intensity.

After the noise is added to Eq. (2), the equation becomes:

$$\begin{cases} \varepsilon \frac{dv}{dt} = v(v-a)(1-v) - w \\ \frac{dw}{dt} = v - dw - b + r\sin\beta t + D \cdot \xi(t) \end{cases} \tag{3}$$

Ref. [2] researched the strange non-chaotic attractor in logistic map and kicked duffing oscillator when the system bifurcation parameter is in the right of boundary crisis. In this case, a periodic attractor and a non-attracting chaotic invariant set (chaotic saddle) coexist, and the asymptotic attractor of system is periodic at the absence of noise. Under the perturbation of appropriate noise, the transition between periodic attractor and chaotic saddle will occur. Then strange non-chaotic attractor is induced. The attractor appears to be geomertrically strange, but its nontrivial maximum Lyapunov exponent is negative. There are fluctuations of the finite-time Lyapunov exponent into the positive side.

According to the Section 2, a boundary crisis occurs in FHN neuron model when b is 0.2404. When the b is in the right of boundary crisis point, a periodic-3 attractor and a non-attracting chaotic invariant set (chaotic saddle) coexist, and system displays periodic-3 spiking at the absence of noise. At the presence of appropriate noise, the same transition between periodic attractor and chaotic saddle will occur.

Then strange non-chaotic attractor should be induced to occur. The mechanism of strange non-chaotic attractor in this case is the same as the case in Ref. [2].

When the b is in the left of boundary crisis, the asymptotic attractor of system is a chaotic attractor. Then, the Lyapunov exponent of system response is positive. According to Ref. [7], at the presence of weak noise, the system motion may be in transition between attractors in two sides of continuous bifurcation point. Appropriate noise can induce the system motion to move towards attractors in the right of crisis point. Trajectory from a random initial condition typically moves toward the chaotic attractor, stays near the attractor for a finite time, and noise induced it to transit to the periodic attractor or chaotic saddle. If the system motion transit to periodic attractor, it will stay finite time in periodic attractor until noise enough to induce it to transit to chaotic saddle or chaotic attractor in the left of crisis point. If the system motion transit to chaotic saddle, there is thus transient chaos. Trajectory typically moves toward the chaotic saddle along its stable manifold, stays near the saddle for a finite time, and leaves the saddle along its unstable manifold before finally approaching the periodic attractor. After finite time, the appropriate noise also moves the system motion to the chaotic attractor in the left of boundary crisis point. Then strange non-chaotic attractor should be also induced to occur. The mechanism of strange non-chaotic attractor is related to the random transitions between chaotic attractor in the left of boundary crisis point and the periodic attractor in the right of crisis point. Because in the right of crisis point there coexist chaotic saddle and periodic attractor, the system motion also transits probably between periodic attractor and chaotic saddle. Because chaotic spike neurons are more sensitive to stimulis than periodic firing neurons [4], strange non-chaotic attractor in this case is of much more significance to reveal some phenomena in life science.

To verify the theoretical analysis above, the Eq. (3) is integrated using the fourth-order Runge-Kutta method with time step $\Delta t = 0.01$ second. And the Poincare maps of several cases are obtained respectively by section $t = n^*2\pi/\beta (n = 1, 2, 3, \ldots)$. The Lyapunov exponent λ_1 of system in several cases are calculated respectively. The results of calculation are shown in Figs. 2, 3, 4, and 5. From the

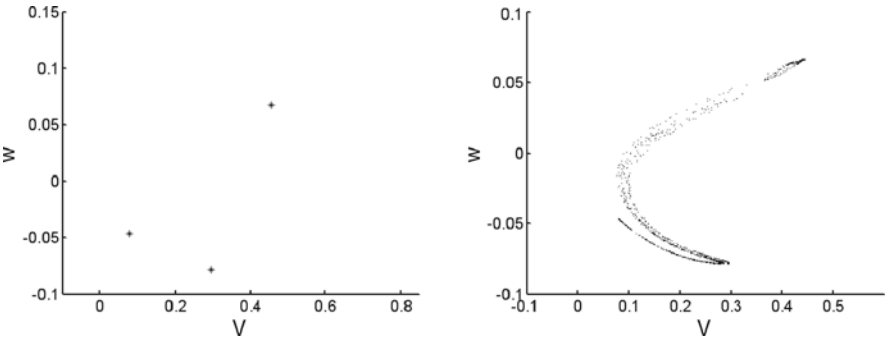


Fig. 2 The attractors of FHN neuron model (Eq. 3) in the *right* of crisis point when $b = 0.241$, $\varepsilon = 0.02$, $d = 0.78$, $a = 0.5$, $r = 0.27$, $\beta = 14.0$. (**a**) $D = 0.0$, (**b**) $D = 0.005$

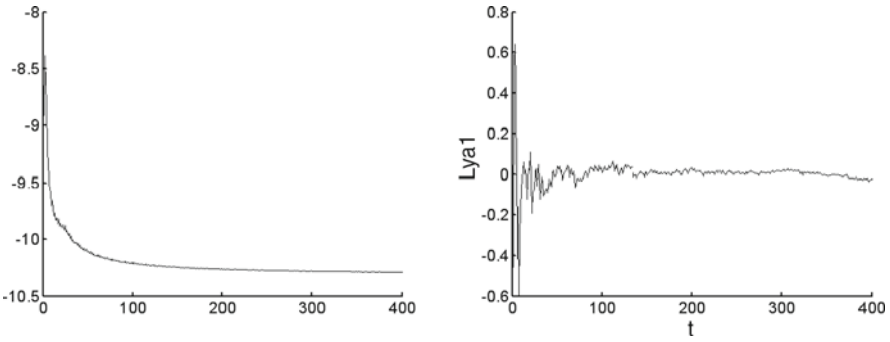


Fig. 3 The evolution of the finite-time Lyapunov exponent in two cases of noise in the *right* of crisis point when $b = 0.241,\ \varepsilon = 0.02,\ d = 0.78,\ a = 0.5,\ r = 0.27,\ \beta = 14.0$. (**a**) D=0.00, (**b**) $D = 0.005$

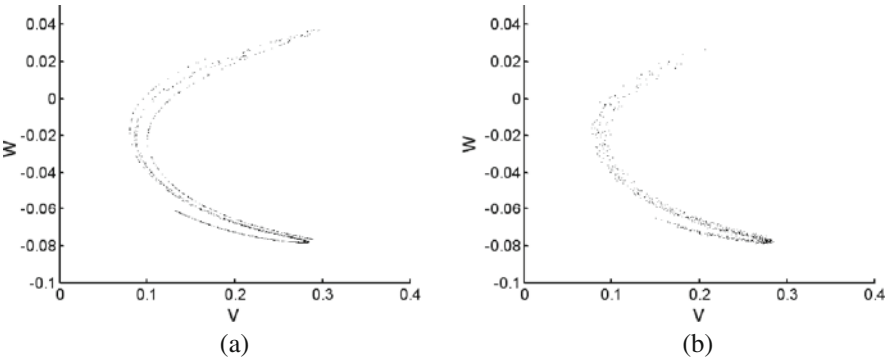


Fig. 4 The attractors of FHN neuron model (Eq. 3) in the *left* of crisis point when $b = 0.240,\ \varepsilon = 0.02,\ d = 0.78,\ a = 0.5,\ r = 0.27,\ \beta = 14.0$. (**a**) $D = 0.0$, (**b**) $D = 0.03$

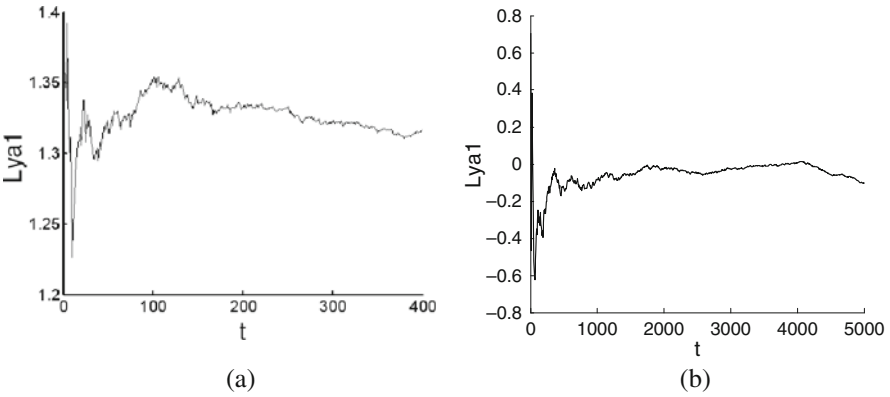


Fig. 5 The evolution of the finite-time Lyapunov exponent in two cases of noise in the *left* of crisis point when $b = 0.240,\ \varepsilon = 0.02,\ d = 0.78,\ a = 0.5,\ r = 0.27,\ \beta = 14.0$. (**a**) $D = 0.00$, (**b**) $D = 0.03$

Figs. 2 and 3, it can be seen that when b is in the right of crisis point the asymptotic attractor is a period-3 attractor without noise, the Lyapunov exponent λ_1 is negative. And at the presence of weak noise the asymptotic attractor is a strange attractor, but it is not chaotic because its Lyapunov exponent is non positive at the most time. From the Fig. 4 and 5, it can be seen that when b is in the left of crisis point the asymptotic attractor is a chaotic attractor without noise, the Lyapunov exponent λ_1 is positive. And at the presence of weak noise the asymptotic attractor is a strange attractor, but it also is not chaotic because its Lyapunov exponent is non positive at the most time.

4 Conclusions

In summary, we have shown that strange non-chaotic attractor can occur in periodically driven FHN neuron. With the change of b system leads to chaos by the route of periodic-doubling bifurcation. When b is some certain value crisis occurs. And when $b = 0.2404$, boundary crisis occurs. A chaotic saddle in FHN neuron model appears after boundary crisis occurs. In a certain parameter region before boundary crisis there coexist a chaotic attractor and a period-3 attractor. But after boundary crisis chaotic attractor disappears, the system displays a period-3 spiking, and a chaotic saddle remains. Not only can strange non-chaotic attractor appear at the appropriate noise when there coexist periodic attractor and chaotic saddle in nonlinear dynamical system just like Ref. [3], but also when the asymptotic attractor of system is a chaotic attractor in the neighborhood of left of crisis point.

Acknowledgments The National Science Foundation under Grant (10872156, 30670660, 10432010), Shaanxi province Science Foundation under 2007014 and National post-doctor Foundation under 20080430203 support the work reported in this paper.

References

1. Gong, Y.F., Xu, J.X.: Stochastic resonance: when does it not occur in neuronal models?. Phys. Lett. A. **243** (1998) 351–359.
2. Grebogi, C., Ott, E., Pelikan, S., et al.: Strange attractors that are not chaotic. Physica D. **13**(1–2) (1984) 261–268.
3. Wang, X., Zhan, M., Lai, C.-H., Lai, Y.-C.: Strange non-chaotic attractors in random dynamical systems. Phys. Rev. Lett., **92**(7) (2004) 047102.
4. Hu, S.J., Yang, H.J., Jian, Z., Long, K.P., Duan, Y.B., Wang, Y.H., Xing, J.L., Xu, H., Ju, G.: Adrenergic sensitivity of neurons with non-periodic firing activity in rat injures dorsal root ganglion. Neuroscience. **101**(3) (2000) 689–698.
5. FitzHugh, R.: In Schwann, H.P. eds. Biological Engineering, New York, NY: McGraw-Hill (1962); Nagumo, J., Arimoto, S., Yoshizawa, S.: Proc. ZRE **50** (1962) 2061.
6. Zou, H.: Global dynamics for nonlinear dynamical systems and unstable invariant sets speciality (in Chinese), Dissertation for Master Doctor Degree, Xi'an Jiaotong University (2007).
7. Zhang, G.J., Xu, J.X.: Characteristic of nonlinear system stochastic resonance in the neighborhood of bifurcation point. *Acta of Physica Sinica* **54**(2) (2005) 557–564 (in Chinese).

Pattern Discrimination on Deterministic and Noise-Induced Impulses in Neural Spike Trains

Ying Du and Qi-Shao Lu

Abstract There are considerable contrasts in between the interval distribution of firing responses caused by purely deterministic simulations and noisy ones. We use ISI-distance to calculate the distance between the deterministic and noise-induced impulse patterns of the cold receptor model. This result is instructive to understand how noise interacts with non-linear dynamics in encoding mechanisms of neuronal stimulus by using data analysis on spike trains extracted from the model.

Keywords Spike train · Noise · ISI-distance

1 Introduction

Noise introduced to nonlinear systems can have a profound effect on their dynamics [1]. Neuronal models are among the most widely used dynamics to study the phenomenology of excitable systems under the influence of noise. When these models are driven by noise, a variety of excitation phenomena including stochastic resonance and coherence resonance has been observed. Especially, signal encoding in temperature sensitive skin receptors is attributed to modulation of intrinsic membrane potential oscillations with contribution of noise [2].

For further examination of the principle encoding properties and the effects of noise, Braun et al. used a minimal Hodgkin-Huxley type computer model to simulates stationary temperature dependencies of peripheral cold receptors [3]. This model shows bifurcations from tonic firing to burst discharges, modulation of burst discharges and transitions from subthreshold oscillations to skipping and regular firing [4]. In this paper, we study the noise effects in the temperature encoding

Y. Du (✉)
The Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237 China; Department of Dynamics and Control, Beihang University, Beijing 100191, China
e-mail: du.ying.buaa@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_51,

of neuronal spike trains in this cold receptor model by using data analysis of ISI-distances of firing patterns caused by purely deterministic simulations and noisy ones.

2 The ISI-Distance Method and Analysis

The corresponding data analysis suggest that addition of noise can considerably extend the dynamical behavior of the system [4]. First, we use the return maps of inter-spike interval(ISI), that is ISI(n+1) versus ISI(n), to reflect the influence of noise shown as Fig. 1.

From the Fig. 1, we can see that there are additional clusters which include both longer intervals indicating the occurrence of skippings as well as very short intervals indicating burst discharges. The noisy return map elucidates obvious different fluctuations with the different noise density. We can see from the figure that the effect of the noise become greater with increasing the noise intensity.

Except traditional methods for analyzing the data, measuring the overall degree of synchrony between different spike trains is also an important tool. In this paper,

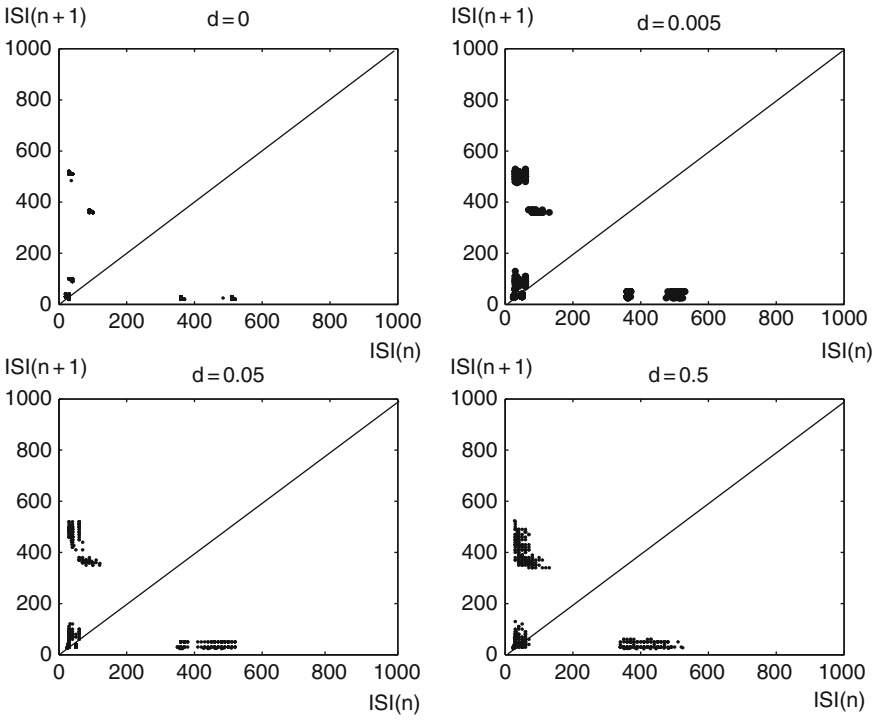


Fig. 1 Return maps of simulation runs for the same temperature $T = 20°C$ but different levels of the noise intensity

we also use ISI-distance [5] to compare dissimilarity between the deterministic and noisy simulations at different temperatures. The distance between two spike trains is defined by the following procedure: the value of the current interspike interval of the spike train is defined according to each time instant

$$x_{isi}(t) = min(t_i^x | t_i^x > t) - max(t_i | t_i^x < t), t_1^x < t < t_M^x \tag{1}$$

and accordingly for the other spike train t_j^y. Then the second step the ratio between two spike trains x_{isi} and y_{isi} is taken

$$I(t) = \begin{cases} \frac{x_{isi}(t)}{y_{isi}(t)} - 1, & if\ if x_{isi}(t) < y_{isi}(t) \\ -(\frac{y_{isi}(t)}{x_{isi}(t)} - 1), \text{ else} \end{cases} \tag{2}$$

Finally, the absolute *ISI*-distance is integrated over time,

$$D_I = \int_{t=0}^{T} dt|I(t)|, \tag{3}$$

Then we use this ISI-distance to characterize the distances between deterministic and noised spike trains. The Fig. 2 presents the firing regimes at $T = 20°C$. The middle row shows the two recorded time series where the spike train x is deterministic data (marked in blue) and y is a stochastic spike train with varying noise level (marked in red).

The ISI-values according to Eq. (1) are shown in each top row, and the corresponding normalized ISI-distance Eq. (2) are shown in the bottom row. The value of the ISI-distance sum appears as a label in the title of each subfigure. Each subfigure shows the distance between the deterministic spike train and the stochastic one at varying levels of the noise intensity. We find as the noise density increases from $d = 0.005$ to $d = 0.5$ in each row, the fluctuations of the spike times in y(t) become stronger. The ISI-distance between the spike train x(t) and y(t) increases form $0.178(d = 0.0005)$ to $0.389(d = 0.05)$, and then to $0.446(d = 0.5)$. That is, the ISI-distance will increase for stronger noise intensity.

3 Conclusion

In conclusion, we found that the effect of the noise will become stronger with increasing the noise intensity, that is, spike trains can be more strongly affected by greater noise than lower one in cold receptors. This again indicates that addition of noise can considerably extend the dynamical behavior of the system with coexistence of different dynamical situations at deterministically fixed parameter. This effect between low- and high-dimensional dynamics have to be considered as qualitatively important factors in neuronal encoding.

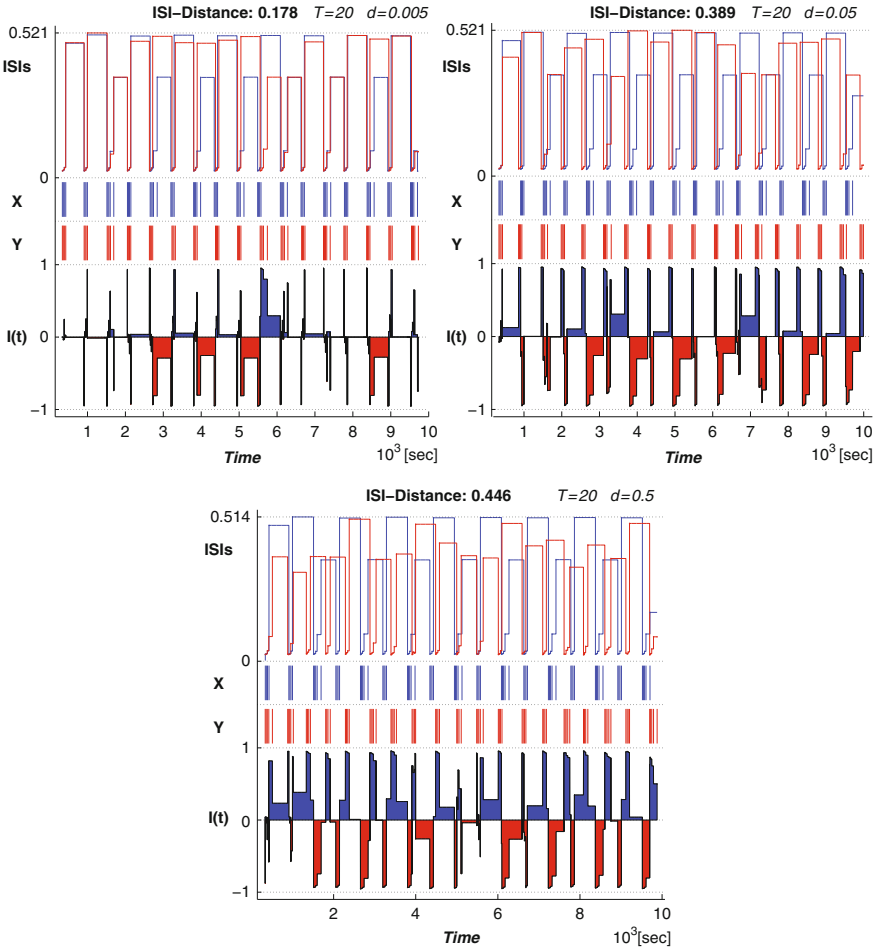


Fig. 2 The ISI-distance between the deterministic and stochastic spike trains at $T = 20°C$ with different noise

Acknowledgement This work was supported by the National Natural Science Foundation of China (No. 10872014,10872068,10672057).

References

1. Horsthemke, W., Lefever R.: Noise-Induced Transitions. Berlin: Springer (1984).
2. Lindner, B., Garcia a Ojalvo, J., Neiman, A.: Schimansky-Geier L. Effects of noise in excitable systems. Phys. Rep. **392** (2004) 321–424.
3. Braun, H.A, Bade, H., Hensel, H.: Static and dynamic discharge patterns of bursting cold fibers related to hypothetical receptor mechanisms. Pflügers. Arch. **386** (1980) 1–9.

4. Braun, H.A., Huber, M.T., Dewald, M., Schäfer, K., Voigt, K.: Computer simulations of neuronal signal transduction: the role of nonlinear dynamics and noise. Int. J. Bifurcat. Chaos. **8** (1998a) 881–889.
5. Kreuz, T., Haas, J.S., Morelli, A., Abarbanel, H.D.I., Politi, A.: Measuring spike train synchrony. J. Neurosci. Meth. **165** (2007) 151–161.

Desynchrony in Two Chaotic Neurons with Uncertain Couple and Mix-Adaptive Feedback

Yong Zhao and Qishao Lu

Abstract In this Chapter, desynchrony in two chaotic neurons with uncertain couple and adaptive feedback is discussed. A new adaptive feedback scheme is proposed to realize desynchrony in two chaotic neurons. At last, numerical simulations for two Hindmarsh-Rose neurons model. Maybe there is a theoretical guide to understand information processing and coding, even treating some diseases.

Keywords desynchrony · Adaptive delay feedback · Neural activity · The synaptic coupling strengths · Neuronal network · The feedback strength

1 Introduction

Neural activity is a cooperative process of neurons [1]. Collective behaviour of neural activity plays an important role in the transmission of the neural information. Synchronization as a Collective behaviour plays a vital role in information processing in the brain. However, physiological studies have shown some diseases have been caused by too much synchronization, such as Parkinsons disease. Therefore, for seeking controlling method of synchronous neurons, maybe there is a theoretical guide to understand information processing and coding, even treating some diseases. Hence, as an example, the uncertain coupled Hindmarsh-Rose neurons model is discussed by the given scheme in Section 3. Furthermore, we find small mismatch parameters on the effects of the desynchrony are robust.

Y. Zhao (✉)
Department of Dynamics and Control, Beihang University, Beijing 100083, China
e-mail: zhaoyong_54@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_52,

2 Main Result

In this section, based on the invariance principle of differential equations [2] and matrix inequality, desynchrony in two uncertain coupled chaotic neurons with adaptive delay feedback is investigated,and general theory is given. Consider the following two uncertain coupled chaotic neurons system:

$$\dot{x}_1 = f(x_1) + \varepsilon H(x_1, x_2),\ dotx_2 = f(x_2) + \varepsilon H(x_2, x_1) \tag{1}$$

where $x_i = (x_{i1}, x_{i2}, \cdots, x_{in})^T \in R^n, f(x_i) = (f_{i1}(x_i), f_{i2}(x_i), \cdots, f_{in}(x_i))^T$ is a nonlinear vector function, H is unknown, and can be linear or nonlinear. Let $\Omega \subset R^n$ be a chaotic bounded set of (1), which is globally attractive. Generally, we assume:

(1) the nonlinear vector function $f(x_i)$ is globally Lipschitz with positive constants $k_i > 0$ such that $|f_i(x) - f_i(y)| \le k_i|x - y|$ for any $x(t),\ y(t) \in R^n$.
(2) $H(x_i, x_j)$ is Lipschitz with positive constants $m_i > 0$ and n_j such that $|H(x_i, x_j) - H(y_i, y_j)| \le m_i|x_i - y_i| + n_j|x_j - y_j|$.

In what follows, two uncertain coupled chaotic neurons with mix-adaptive feedback is described by the following equation.

$$\begin{cases} \dot{y}_1 = f(y_1) + \varepsilon H(y_1, y_2) + k_1(x_1 - y_1), \\ \dot{y}_2 = f(y_2) + \varepsilon H(y_2, y_1) + k_2(x_2(t - \tau) - y_2) \end{cases} \tag{2}$$

where y is the response state, τ is the delay for the feedback control. The feedback strength $k_i = (k_{i1}, k_{i2}, \cdots, k_{in})^T$ is adapted according to the following scheme:

$$\dot{k}_1 = \gamma_1(x_1 - y_1)^2,\ \ \dot{k}_2 = \gamma_2(x_2(t - \tau) - y_2)^2 \tag{3}$$

Let $e_i = x_i - y_i,\ e_i^\tau = x_i(t - \tau) - y_i,\ x_i(t - \tau) = x_i^\tau(t)$, then we can obtain from (2.1) and (2.2)

$$\begin{cases} \dot{e}_1 = f(x_1(t)) - f(y_1) + \varepsilon(H(x_1(t), x_2^\tau(t)) - H(y_1, y_2)) + k_1 e_1, \\ \dot{e}_2^\tau = f(x_2^\tau(t)) - f(y_1) + \varepsilon(H(x_2^\tau(t), x_1^\tau(t)) - H(y_2, y_1)) + k_2 e_2^\tau, \end{cases} \tag{4}$$

Definition 1 For the two systems described by Eqs. (1) and (2), we say they possess the property of desynchrony if there exists $e_i \to 0$ and $e_j^\tau \to 0,\ i \ne j,\ i, j = 1, 2$, as $t \to +\infty$.

Theorem *Assume (1) and (2) holds, furthermore, Eqs. (1) and (2) satisfy the adaptive scheme (3). Then desynchrony in two uncertain coupled chaotic neurons can be achieved, that is,* $\lim_{t\to\infty} e_i = 0$ *and* $\lim_{t\to\infty} e_j^\tau = 0, i \ne j,\ i, j = 1, 2.$

According to above proposed desynchrony scheme, we obtain desynchrony in two uncertain coupled chaotic neurons effectively. The above results are rather general. In what follows, the numerical example may be helpful to understand the above analytical results.

3 Numerical Simulation

In this section, we consider a model Hindmarsh-Rose neuron described by the following equations :

$$\begin{cases} \dot{x}_{i1} = x_{i2} - ax_{i1}^3 + bx_{i1}^2 - x_{i3} + I_x, \\ \dot{x}_{i2} = c - dx_{i1}^2 - x_{i2}, \\ \dot{x}_{i3} = r[s(x_{i1} - \bar{x}) - x_{i3}] \end{cases} \tag{5}$$

where $x_i = (x_{i1}, x_{i2}, x_{i3})$, $i = 1, 2$, x_{i1} is the membrane potential, x_{i2} is a recovery variable associated with fast current, z is a slow ionic current. Here $a = 1.0$, $b = 3.0$, $c = 1.0$, $d = 5.0$, $s = 4.0$, $\bar{x} = -1.60$, $r = 0.005$. I as a control parameter. The change of the value of I causes rich firing behaviour. System(5) exhibits multiple time-scale chaotic bursting behavior for $1.47 < I < 3.45$.

Consider the following two uncertain coupled HR chaotic neurons system:

$$\begin{cases} \dot{x}_{11} = x_{12} - ax_{11}^3 + bx_{11}^2 - x_{13} + I_x + \varepsilon H(x_{21}, x_{11}), \\ \dot{x}_{12} = c - dx_{11}^2 - x_{12} + \varepsilon H(x_{22}, x_{12}), \\ \dot{x}_{13} = r[s(x_{11} - \bar{x}) - x_{13}] + \varepsilon H(x_{23}, x_{13}), \\ \dot{x}_{21} = x_{22} - ax_{21}^3 + bx_{21}^2 - x_{23} + I_x + \varepsilon H(x_{11}, x_{21}), \\ \dot{x}_{22} = c - dx_{21}^2 - x_{22} + \varepsilon H(x_{12}, x_{22}), \\ \dot{x}_{23} = r[s(x_{21} - \bar{x}) - x_{23}] + \varepsilon H(x_{13}, x_{23}), \end{cases} \tag{6}$$

where relevant parameter values are the same of system (5), when $\varepsilon = 1.4$, synchronization has been achieved, the corresponding numerical results are shown in Fig. 1.

As real coupled neurons are concerned, information from the presynaptic neuron is transmitted to the postsynaptic one with certain time delay. Time delays are inherent in neuronal transmissions. Thus it is important to seek controlling method of synchronous neurons.

In what follows, the corresponding systems with adaptive feedback can be described by the following equation:

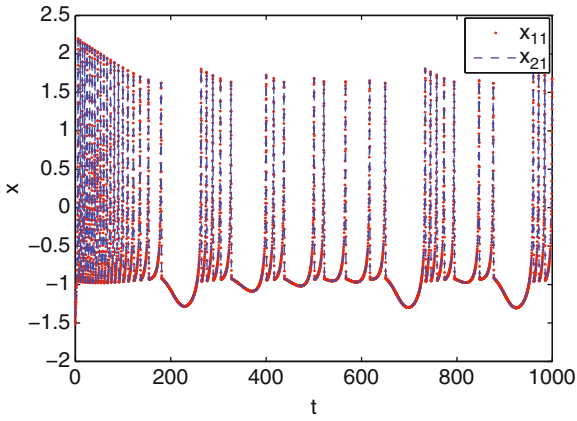


Fig. 1 The portrait of system (3.2) for $\varepsilon = 1.4$

$$\begin{cases} \dot{y}_{11} = y_{12} - ay_{11}^3 + by_{11}^2 - y_{13} + I_y + \varepsilon H(y_{21}, y_{11}) + k_{11}e_{11}, \\ \dot{y}_{12} = c - dy_{11}^2 - y_{12} + \varepsilon H(y_{22}, y_{12}) + k_{12}e_{12}, \\ \dot{y}_{13} = r[s(y_{11} - \bar{y}) - y_{13}] + \varepsilon H(y_{23}, y_{13}) + k_{13}e_{13}, \\ \dot{y}_{21} = y_{22} - ay_{21}^3 + by_{21}^2 - y_{23} + I_y + \varepsilon H(y_{11}, y_{21}) + k_{21}e_{21}^{\tau}, \\ \dot{y}_{22} = c - dy_{21}^2 - y_{22} + \varepsilon H(y_{12}, y_{22}) + k_{22}e_{22}^{\tau}, \\ \dot{y}_{23} = r[s(y_{21} - \bar{y}) - y_{23}] + \varepsilon H(y_{13}, y_{23}) + k_{23}e_{23}^{\tau}, \end{cases} \tag{7}$$

with the law (3) for the adaptive feedback strength, τ represents the time delay.

We take the synaptic coupling strength as k_{i1} and let $k_{i2} = k_{i3} = 0$, $i = 1, 2$. Furthermore, we take the parameters $\gamma_1 = \gamma_2 = 0.1$, $I_x = 3.2$ and the initial feedback strength 0.01 to investigate desynchrony in two uncertain coupled chaotic neurons. This coupled system obviously satisfies the theoretical conditions of the theorem in Section 2 [3]. The corresponding numerical results are shown in Fig. 2. It is seen from Fig. 2a, as $\tau = 0.3$, the coupled system achieves desynchrony. Furthermore, from Fig. 2b, the feedback strength $k_{11} = k_{21}$ asymptotically converges to a local fixed value.

Nonidentity is common in nature. Thus it is important to investigate desynchrony in two uncertain coupled chaotic neurons with small mismatch of parameters. We take $I_x = 3.2$, $I_y = 3.18$, respectively. We can observe the robustness of feedback strength of deynchrony with respect to the feedback strength, compared Fig. 2. This paper proposes a general method for determining desynchrony condition of neuronal networks.

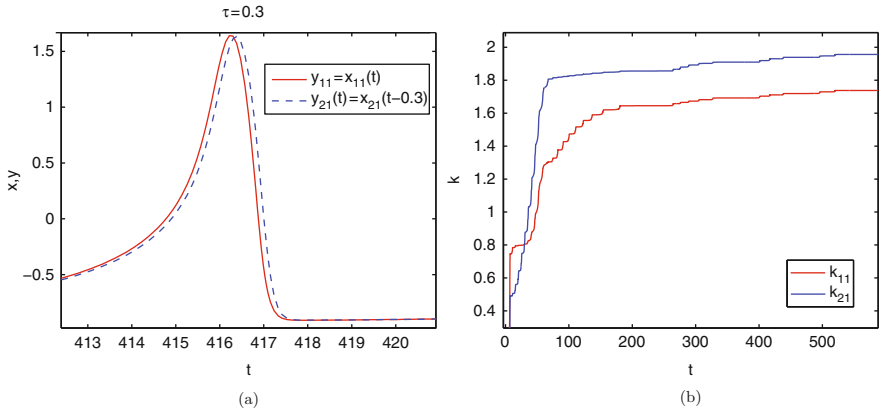


Fig. 2 **a** Desynchrony for two uncertain coupled chaotic neurons with delay $\tau = 0.3$. **b** The temporal evolution of k_{11}, k_{21} for $I_x = 3.2$

Acknowledgements This work was supported by the National Natural Science Foundation of China (No.10872014).

References

1. Gray C.M., König P., Engel A.K., Singer W.: Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature **338** (1989) 334–337.
2. Lasalle J.P.: Some extensions of Lyapunov's second method. IEEE. Trans. Circuit. Theory. **7** (1960) 520–527.
3. Wang Q.Y., Lu Q.S., Duan Z.S.: Adaptive lag synchronization in coupled chaotic systems with unidirectional delay feedback. Int. J. Non. Linear. Mech. doi:10.1016/j.ijnonlinmec.2009.01.001.

Outer Synchronization of Coupled Discrete-Time Networks with Time Delay

Weigang Sun, Rubin Wang, Weixiang Wang, and Jianting Cao

Abstract In this Chapter, we study the outer synchronization between two discrete-time networks with time delay. Firstly we introduce time delay into our models and discuss the delay effects. The synchronization conditions are derived. Finally a numerical example is shown to illustrate our theoretical results.

Keywords Complex network · Inner synchronization · Outer synchronization

1 Introduction

For recent ten years, complex networks have attracted much attention. Network model and network dynamics are two important topics in the studying complex networks. For network models, the small-world network [1] and scale-free network [2] are the most known models. While network synchronization, robustness and fragility of network, virus spreading etc are main research contents in network dynamics.

While studying the dynamics of networks, synchronizing all the nodes is an interesting thing and has attracted wide attention. From the existing papers, we find that synchronization also happens inside a network, strictly speaking, we call it as inner synchronization [3, 4]. A nature question: does synchronization between two coupled networks also happen? In the literature [3], the authors used the master-slave method to realize the synchronization between two coupled networks, proposing the outer synchronization concept. In Ref. [5], the authors study the synchronization between two groups, aiming at the connection between two groups. In natural world, the inner and outer synchronization may be happen at the same time. For example,

W. Sun (✉)
Institute for Cognitive Neurodynamics, School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China; The School of Science, Hangzhou Dianzi University, Hangzhou 310018, China
e-mail: wgsun@hdu.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_53,

in Shanghai subway, when the train reaches the station, the outer and inner doors opened and closed simultaneously.

In the present paper, we study the synchronization between two discrete-time networks with time delay. Due to the distance between two networks, the introduction of time-delay is meaningful. The rest of this paper is organized as follows. In Section 2, the model presentation and synchronization analysis is given, numerical example is shown in Section 3, finally the conclusion is presented.

2 Model Presentation and Synchronization Analysis

We consider the following network models, the node dynamical equation is $x_i(t+1) = f(x_i(t)), i = 1, \cdots, N_x$ and $y_j(t+1) = g(y_j(t)),\ j = 1, \cdots, N_y$, respectively, where $f(\cdot) : R^{n_x} \to R^{n_x}, g(\cdot) : R^{n_y} \to R^{n_y}$ are continuously differential functions, $x_i(y_j)$ is an n_x-dimensional (n_y-dimensional) state vector. The dynamical equations of the network systems are as follows:

$$\begin{aligned} x_i(t+1) &= f(x_i(t)) + \sum_{j=1}^{N_y} a_{ij}\Gamma_1 y_j(t-\tau), i = 1, 2, \cdots, N_x, \\ y_j(t+1) &= g(y_j(t)) + \sum_{i=1}^{N_x} b_{ji}\Gamma_2 x_i(t-\tau), j = 1, 2, \cdots, N_y, \end{aligned} \tag{1}$$

where A is an $N_x \times N_y$ dimensional coupling matrix, whose entries (a_{ij}) represent the intensity of the direct interaction from i in network X to j in network Y, analogously the entries of (b_{ji}) are same defined as (a_{ij}). Matrix $\Gamma_1(\Gamma_2) \in R^{n_x \times n_y}(R^{n_y \times n_x})$ is the inner-coupling matrix. τ is time delay between networks.

Let's now consider the possibility whether the individual network achieve synchronization, i.e., $x_1(t) = \cdots = x_{N_x}(t) = x_s(t)$ and $y_1(t) = \cdots = y_{N_y}(t) = y_s(t)$. If there exist such synchronous states, satisfying

$$\sum_{j=1}^{N_y} a_{ij} = \mu_1, \forall i \in X, \sum_{i=1}^{N_x} b_{ji} = \mu_2, \forall j \in Y.$$

without loss of generality, we set $\mu_1 = \mu_2 = 1$.

Thus the synchronized state equations are

$$\begin{aligned} x_s(t+1) &= f(x_s) + \Gamma_1 y_s(t-\tau), \\ y_s(t+1) &= g(y_s) + \Gamma_2 x_s(t-\tau). \end{aligned} \tag{2}$$

Linerizing the synchronous state around x_s and y_s, we get

$$\begin{aligned} \delta x_i(t+1) &= Df(x_s)\delta x_i(t) + \sum_{j=1}^{N_y} a_{ij}\Gamma_1 \delta y_j(t-\tau), i = 1, 2, \cdots, N_x, \\ \delta y_j(t+1) &= Dg(y_s)\delta y_j(t) + \sum_{i=1}^{N_x} b_{ji}\Gamma_2 \delta x_i(t-\tau), j = 1, 2, \cdots, N_y. \end{aligned} \tag{3}$$

Applying the algebraic knowledge, we derive the following equations,

$$\begin{aligned} \delta\bar{x}(t+1) &= Df(x_s)\delta\bar{x}(t) + \lambda\Gamma_1\delta\bar{y}(t-\tau), \\ \delta\bar{y}(t+1) &= Dg(y_s)\delta\bar{y}(t) + \lambda\Gamma_2\delta\bar{x}(t-\tau), \end{aligned} \tag{4}$$

where λ is the eigenvalues of matrix $\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$.

In the sequel, we utilize the linear matrix equality to drive a synchronous theorem.

Theorem 1 *Consider network model (1). If there exist two positive matrices* $P, Q > 0$, *satisfying* $\Theta =$

$$\begin{pmatrix} \Phi_1 & 0 & 0 & \lambda Df(x_s)^T P\Gamma_1 \\ 0 & \Phi_2 & \lambda Dg(y_s)^T Q\Gamma_2 & 0 \\ 0 & \lambda\Gamma_2^T QDg(y_s) & \lambda^2\Gamma_2^T Q\Gamma_2 - I_{n_x} & 0 \\ \lambda\Gamma_1^T PDf(x_s) & 0 & 0 & \lambda^2\Gamma_1^T P\Gamma_1 - I_{n_y} \end{pmatrix} < 0,$$

where $\Phi_1 = Df(x_s)^T PDf(x_s) - P + I_{n_x}, \Phi_2 = Dg(y_s)^T QDg(y_s) - Q + I_{n_y}$, *and* I *denotes the identity matrix, then the network (1) asymptotically synchronizes to* x_s, y_s *defined by the equation (2) for the fixed delay* $\tau > 0$, *respectively.*

Proof Consider the Eq. (4). Choose the Lyapunov function as

$$\begin{aligned} V(t) &= \delta\bar{x}(t)^T P\delta\bar{x}(t) + \delta\bar{y}(t)^T Q\delta\bar{y}(t) \\ &\quad + \textstyle\sum_{\sigma=t-\tau}^{t-1} \delta\bar{x}(\sigma)^T\delta\bar{x}(\sigma) + \sum_{\sigma=t-\tau}^{t-1} \delta\bar{y}(\sigma)^T\delta\bar{y}(\sigma). \end{aligned} \tag{5}$$

Therefore,

$$\Delta V(t) = V(t+1) - V(t) = \begin{pmatrix} \delta\bar{x}(t) \\ \delta\bar{y}(t) \\ \delta\bar{x}(t-\tau) \\ \delta\bar{y}(t-\tau) \end{pmatrix}^T \Theta \begin{pmatrix} \delta\bar{x}(t) \\ \delta\bar{y}(t) \\ \delta\bar{x}(t-\tau) \\ \delta\bar{y}(t-\tau) \end{pmatrix}. \tag{6}$$

From the above condition, we know the zero solution of (4) is asymptotically stable, which shows that the networks achieve individual synchronization.

3 Numerical Example

In the networks considered below, the node dynamics are taken as the Logistic map, i.e., $f(x) = \rho_1 x(1-x)$ and $g(y) = \rho_2 y(1-y)$.

The coupling matrices are chosen as,

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 1/4 & 1/4 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 3/4 & 0 \end{pmatrix}.$$

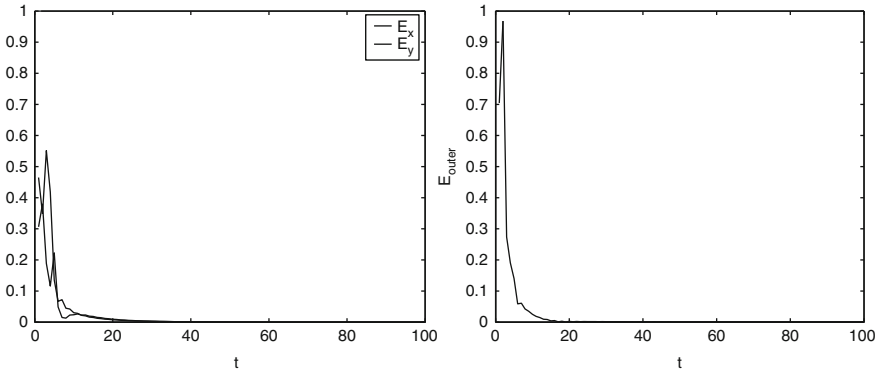


Fig. 1 Synchronization errors of network (1) with $\rho_1 = 2.6, \rho_2 = 2.3, \tau = 1$

We take Γ_1, Γ_2 as identity matrices of dimension one. Set $E_x = \|x_i - x_s||, i = 1, \cdots, 3$, and $E_y = \|y_j - y_s||, j = 1, \cdots, 3$, and $E_{\text{outer}} = \|x_i - y_i||, i = 1, \cdots, 3$. Figure 1 plots the synchronization errors.

Taking some fit parameters, we find the networks achieve the individual synchronization, and outer synchronization also exists. This is interesting. Especially the delay influences the outer synchronization more.

4 Conclusion

In this paper, we study the synchronization between two discrete-time networks with time delay. From the numerical example we can see the delay affects their inner network synchronization less, while the delay has greater influence on outer synchronization. How to give the domain of time delay will be the scope of our future research work.

References

1. Watts, D.J., Strogatz, S.H.: Collective dynamics of small-world networks. Nature **393** (1998) 440–442.
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286** (1999) 509–512.
3. Li, C.P., Sun, W.G., Kurths, J.: Synchronization between two coupled complex networks. Phys. Rev. E. **76** (2007) 046204.
4. Li, C.P., Sun, W.G., Kurths, J.: Synchronization of complex dynamical networks with time delays. Physica. A. **361** (2006) 24–34.
5. Sorrentino, F., Ott, E.: Network synchronization of groups. Phys. Rev. E. **76** (2007) 056114.

Part VI
Brain Imaging, EEG, MEG

Phase Resetting on Scalp EEG, Study on Its Synchrony and Distribution

Yusely Ruiz, Susan Pockett, Walter J. Freeman, Eduardo González, and Guang Li

Abstract A 64 electrode cap was used to record the EEG from 6 subjects. The Hilbert transform was used to estimate the analytic amplitude and analytic frequency. Phase resetting were identified by spikes on the temporal differences of successive unwrapped analytic phase value. Evidence was acquired of global phase resetting and re-synchronization that is necessary for the spatial-temporal pattern emergence. Phase resetting were clustered in time and this is evidence that the channels are not completely independent. The repetition rates of the phase resetting were proportional to the filter bandwidth by a factor close to 0.641 as predicted by Rice.

Keywords Scalp EEG · Hilbert transform · Phase resetting

1 Introduction

The brain is a complex system, integrated by a variety of functional entities connected at numerous levels: microscopic, mesoscopic and macroscopic. The techniques used to record brain activity at microscopic (local field potentials) and mesoscopic (ECoG) scales are invasive procedures commonly used with animals and occasionally with humans who have brain diseases that require surgery or are victims of global paralysis. However the techniques used at macroscopic scales (EEG, MEG, fMRI), are completely non-invasive procedures. Thus, it is a research interest to study the brain activity at macroscopic scales, taking advantage of the cognitive and phenomenological awareness of a normal healthy subject by non-invasive procedures.

Well-defined behavioral states require sequential epochs of cortical activity, in which the neural activity in multiple cortical areas is formed by cooperation and

G. Li (✉)
National Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, P.R. China
e-mail: guangli@zju.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_54,

expressed in spatial-temporal patterns [1, 2]. Thus far there is some experimental evidence that supports the idea that the sequence formation of spatial–temporal patterns begins with the abrupt resetting of phase values on every channel, followed by re-synchronization and spatial pattern stabilization [3, 4]. Freeman and coworkers had found in the Hilbert Transform (HT) a useful tool to study these patterns, mainly in ECoG signals [3–7]. They have also demonstrated that in human and animal ECoG the distribution and rate of the phase resetting showed properties similar to a "Rice distribution" [8, 9]. In this work the HT is applied to EEG signals and phase resetting synchrony and its distribution is studied.

2 Materials and Methods

Data were collected in the Psychology Department of the University of California Berkeley and the study was approved by the UC Berkeley Institutional Review Board. Data from six healthy male subjects between 23 to 44 years old were recorded and all subjects gave informed consent. The EEG was recorded using BioSemiTM amplifiers with a 64-electrode cap using Ag/AgCl electrodes; the sampling rate was 512 Hz and the analog filter was set from DC to 134 Hz. Continuous records were taken, with the times of various sorts of stimuli and responses marked in a 65th recording channel, more detail about signal recording can by found in [10].

Specific data epochs, surrounding the stimulus mark, were extracted from the recording. The offset of each channel was removed and each epoch was normalized by dividing by the global standard deviation. Temporal band pass filters were then applied to extract beta (12–25 Hz) and gamma (25–50 Hz) and finally the HT was applied to decompose the EEG into the analytic amplitude and analytic phase [5, 7].

Synchrony of phase resetting across the all array was explored by naked eye using the CAPD raster plot (Coordinate Analytic Phase Differences defined as the temporal differences between the unwrapped analytic phases) [6, 7]. Upward or downward deviation from the mean differences corresponds to phase resetting and they were visualized as spikes. In order to avoid common reference artifactual synchrony, data were re-referenced to the average signal of each hemisphere [10, 11].

The distribution and rate of the phase resetting were also studied using 8 different temporal filters with bandwidth varying from 4 to 32 Hz and centered at two frequencies. Histogram of the time lapse between phase resetting were constructed and the modal frequency, in Hz, was estimated as the inverse of the modal interval, in ms, which was determined from the histogram maxima [9].

3 Results

Abrupt phase resetting were observed involving groups of channel clustered in time but not simultaneous in the original signal and the re-referenced signal in all subject, epoch and frequency bands. Figure 1 show an example of the CAPD raster plot. The

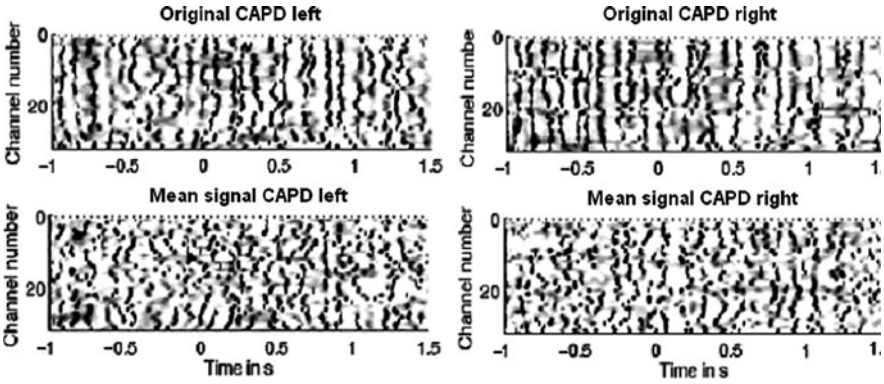


Fig. 1 Raster plot of CAPD for the signal at the original reference and the signal reference to mean signal of each hemisphere. Spikes on CAPD (*black*) formed cluster of electrode that happen at similar times for both references. Zero time present the stimulus onset, Beta band

clustering is evidence that the channels are not independent. Also, phase resetting indicates the possibility of phase discontinuities that are necessary for the emergence of spatial patterns that might be related to cognition.

Histogram and modal frequencies of the time between phase resetting in the EEG signal are showed in Fig. 2 for two central frequencies fixed at 16 or 64. The

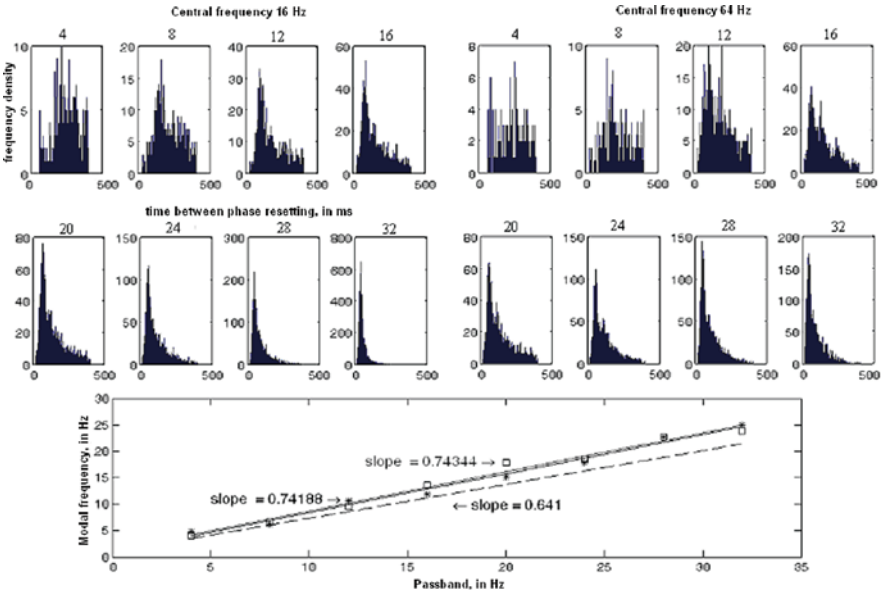


Fig. 2 Histograms of time intervals between phase resetting. The number of spikes and modal frequency is highly dependent of the pass band width. "∗" central frequency at 16 Hz, "□" central frequency at 64 Hz. The numbers on the top of the histograms represent the width of the pass band used

histograms were likewise invariant in accordance with the center frequencies of the pass bands. On the EEG signals the repetition rates of the phase resetting were also proportional to the filter bandwidth, whit small deviations from the 0.641 value at bandwidth wider than 20 Hz. Thus far, it had been argued that these deviations could be introduced by the non-ideal frequency response of the FIR filters [9].

4 Conclusions

Evidence was acquired of global phase resettings and re-synchronization that is necessary for the spatial–temporal pattern emergence. Phase resetting were clustered in time and this is evidence that the channels are not completely independent. Even so referencing to the average signal has the inconvenient consequence of removing a part of the genuine synchrony common to all channels, clusters of phase resetting were also observed in this signal.

The phase resetting rate of EEG signals was proportional to the width of the pass band used as predicted by Rice [8] and demonstrated by Freeman [9] in ECoG signals. This is important because the Rice statistics may provide parameters with which to search for and identify behaviorally related spatial-temporal patterns and also help to understand the mechanisms by which the spontaneous background activity of the brain is generated and reorganized in patterns.

Acknowledgements This research is supported by the Natural Science Foundation of China (Grants No. 60874098 and 60911130129), the National High Technology Research and Development Program of China (863 Program 2007AA042103) and the National Creative Research Groups Science Foundation of China (NCRGSFC: 60421002).

References

1. Fingelkurts, A., Fingelkurts, A.: Timing in cognition and EEG brain dynamics: discreteness versus continuity. Cogn. Process. **7**(3) (2006) 135–162.
2. Freeman, W.J.: Proposed cortical “shutter” mechanism in cinematographic perception. In Neurodynamics of Cognition and Consciousness. New York, NY: Springer (2007) 11–38.
3. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. Clin. Neurophysiol. **116**(5) (2005) 1118–1129.
4. Freeman, W.J.: A cinematographic hypothesis of cortical dynamics in perception. Int. J. Psychophysiol. **60**(2) (2006) 149–161.
5. Freeman, W.J.: Definitions of state variables and state space for brain–computer interface. Part 2. Extraction and classification of feature vectors. Cogn. Neurodyn. **1**(2) (2007) 85–96.
6. Freeman, W.J., Burke, B.C., Holmes, M.D.: Aperiodic phase re-setting in scalp EEG of beta-gamma oscillations by state transitions at alpha-theta rates. Hum. Brain Mapp. **19**(4) (2003) 248–272.
7. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin. Neurophysiol. **115**(9) (2004) 2077–2088.
8. Rice, S.O.: Mathematical analysis of random noise. Bell Sys. Tech. J. **23** (1944) 282–332.
9. Freeman, W.J.: Deep analysis of perception through dynamic structures that emerge in cortical activity from self-regulated noise. Cogn. Neurodyn.

10. Pockett, S., Bold, G.E.J., Freeman, W.J.: EEG synchrony during a perceptual-cognitive task: widespread phase synchrony at all frequencies. Clin. Neurophysiol. in press.
11. Fein, G., Raz, J., Brown, F.F., et al.: Common reference coherence data are confounded by power and phase effects. Electroencephalogr. Clin. Neurophysiol. **69**(6) (1988) 581–584.

The Interaction Between the Parietal and Motor Areas in Dynamic Imagery Manipulation: An fMRI Study

Takafumi Sasaoka, Hiroaki Mizuhara, and Toshio Inui

Abstract Mental imagery is a cognitive function that includes sub-functions such as generation, transformation, and matching. However, the neural substrates for each sub-function are not yet clear. In the present study, we used event-related functional MRI during a modified version of a mental clock task to investigate these neural substrates. While participants were mentally transforming the clock hands, we found activations in the left inferior parietal lobule, left motor related regions (premotor area and supplementary motor area), and left insula, which were contra-lateral to the right hand used to manipulate a 3-D mouse in the learning phase. These results suggest that motor imagery was utilized for transformation of mental imagery.

Keywords Mental imagery · Embodied cognition · Mental rotation · fMRI

1 Introduction

Mental imagery is a fundamental function used for various cognitive processes, such as visual recognition, language processing, and reasoning. This function comprises sub-functions including generation, transformation, and matching. However, the neural substrates corresponding to each sub-function are currently unclear.

The task most frequently used to investigate mental imagery is the mental rotation task. A number of functional imaging studies on mental rotation have reported activation in the posterior parietal cortex (PPC) and in motor related areas. The PPC is suggested to be involved in spatial processing (e.g. [1]). On the other hand, activation in motor related areas is suggested to reflect the use of motor imagery, that is, participants rotate objects mentally using their hands [2]. However, the exact role

T. Sasaoka (✉)
Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
e-mail: sasaoka@i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_55,

of the left and right PPC and the degree of interaction between the PPC and motor related areas in mental rotation tasks are still under debate.

Sack et al. used a mental clock task in which participants generated visual images of two clock faces from auditory presented times (e.g. “2:00” and “5:00”) and compared them [3]. These researchers showed that the left PPC is involved in generation of mental imagery, whereas the right PPC is involved in matching. However, the possibility that other brain regions, e.g. motor cortex, were activated in their task was not discounted.

In the present study, we employed a modified version of the mental clock task and added a component of transformation, in order to clarify the neural substrates for each function involved in mental imagery.

2 Materials and Methods

Twenty-three healthy males, aged 20–37 and all right-handed, participated in the experiment. A written informed consent was obtained from each participant in accordance with the Declaration of Helsinki. The experimental procedure was approved by the ethics committee at Kyoto University.

The experiment consisted of practice and fMRI sessions. In the practice session, participants actively learned the rotation speed of the clock hands (learning phase), and subsequently performed the practice trials (practice phase) outside the MRI scanner.

In the learning phase, one clock face was presented on the CRT display for 5 min. Participants could rotate the clock hands by manipulating a 3-D mouse (3Dconnexion SpaceExplorer) using their right hand, and learned the rotation speed of the clock hands. The clock hands rotated at a constant speed (60°/s) irrespective of the extent of rotation of the 3-D mouse.

In the practice phase, one clock face (1st stimulus) was presented for 1.5 s. Participants were asked to memorize the position of the clock hands. A fixation cross, colored white or red, was presented for 2 s. The color of the fixation cross was used as a cue to indicate the direction of rotation: white denoted clockwise, whereas red denoted counter-clockwise. Following the fixation cross, a dot of the same color as the fixation cross was presented. Participants were asked to imagine rotating the clock hands from the initial position toward the cued direction at the learned speed until the dot disappeared. The duration of the presentation of the dot was varied at 1, 3, or 5 s; that is, the degrees of rotation were 60°, 180°, or 300°, respectively. Following a 1.5 s blank screen, a second clock face (2nd stimulus) was presented. Participants were asked to indicate whether the 2nd stimulus was identical to the imagined position of the clock hands, by pressing the corresponding key using their left forefinger or middle finger. They were asked to respond as quickly and accurately as possible. A beep sound was presented when the response was incorrect. One practice session consisted of 24 trials. Participants completed the learning and practice phases 5 times.

In the fMRI session, participants performed the same task as in the practice sessions. Participants each completed two fMRI sessions. Each session consisted of 72 trials. The fMRI data were collected with a 1.5T MRI scanner (Shimadzu-Marconi) with following scan parameters: TR = 3 s, TE = 49 ms, FA = 90°, FOV = 192 mm, in-plane resolution = 64 × 64, and slice thickness = 5 mm without gap. Image processing and statistical analyses were performed using SPM5.

In order to dissociate the sub-processes in mental imagery, we created regressors for each task event. These were: (1) the period from the onset of the first stimulus until the fixation cross disappeared (Encoding); (2) the presentation of a dot (Transformation); (3) the period from the presentation of the second stimulus until the key response (Matching); and (4) the period from 1 s after the key response until the onset of the 1st stimulus (Baseline). The fMRI data for each event were compared with the baseline.

3 Results

The data from seven participants were excluded because the hit rate was below chance even in the final training session. The rest of the participants self-reported that they did not adopt any alternative strategy without making visual images to perform the task. Figure 1 shows the mean hit rates. Three-way analysis of variance (ANOVA) with factors of angle of rotation, direction, and session revealed significant main effects of angle ($F(2, 30) = 24.39, p < 0.001$) and direction ($F(2, 30) = 6.99$, $p < 0.05$). There were no significant interactions. A post hoc comparison by Tukey's HSD revealed significant differences between all conditions of rotation angle ($p < 0.05$).

Figure 2a shows the brain activation during transformation (uncorrected $p < 0.005$, cluster size > 20 voxels). We found significant activation in the left dorsal premotor cortex (PMd) (BA6), the left supplementary motor area (SMA) (BA6), the left insula (BA13), and the left inferior parietal lobule (IPL) (BA40). Moreover, we found activation that was dependent on the angle of rotation during transformation in the visual areas (BA18), the right fusiform gyrus, and the right occipito-parietal junction (BA7/19) including the right transverse occipital sulcus

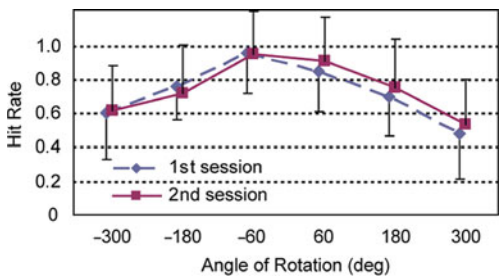


Fig. 1 Hit rate

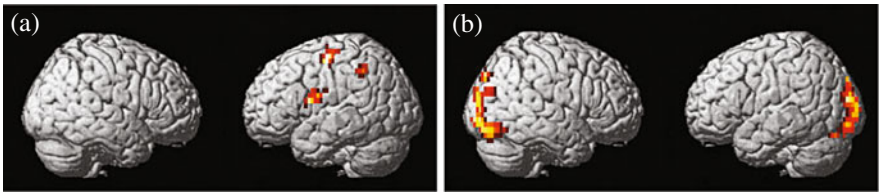


Fig. 2 fMRI results (uncorrected $p < 0.005$, cluster size > 20 voxels) **a** Activation during transformation, **b** angle-dependent activation during transformation

(Fig. 2b). During matching, the pre-SMA (BA6/8), the right ventral premortor cortex (PMv) (BA6/9/44), the right postcentral gyrus (BA2/5), and the right IPL (BA40) were activated. During encoding, the left PM (BA6) and the left superior parietal lobule (SPL) (BA7) were activated.

4 Discussion

During transformation, motor related regions were activated only in the left hemisphere, which is the hemisphere that is contra-lateral to the right hand used to manipulate the 3-D mouse in the learning phase. Kosslyn et al. [2] reported that left motor related regions were activated when participants performed a mental rotation task that involved imagining rotation as a consequence of a motor activity performed by their right hand [2]. This is consistent with our results.

The left PPC is suggested to be involved in generating mental images [3]. Moreover, it has been suggested that the left PPC has an important role in prediction of movement of an object that is caused by self-generated movements [4]. Therefore, the brain activation during the transformation reflects the updating of the position of imagined clock hands in the left IPL by utilizing motor commands from the motor related regions. In addition, during encoding, the left PM was activated. This may reflect the possibility that participants had already started to rotate the clock hands prior to the presentation of the dot.

We found an angle-dependent activation in the right parieto-occipital region. Spatial relationships of visual features are most likely processed in this region, including the right transverse occipital sulcus [5]. The activations in these areas can therefore be interpreted as reflecting processes involved in the calculation of the difference between the first stimulus and the imagined clock face. These activations in the right hemisphere were observed during matching. Activation in the right PPC is consistent with the results of Sack et al. [3]. Ogawa et al. also showed that the right PPC is involved in evaluating visuo-motor prediction error [6]. Therefore, the activations in the right hemisphere during matching might reflect a comparison process between the second clock face and the prediction of a clock face generated in the left parietal and motor areas.

Acknowledgement This study was supported by Grant-in Aid for Scientific Research (S) No. 20220003 from the Japanese Ministry of Education, Science, Technology, Sports and Culture (MEXT).

References

1. Cohen, M.S., et al.: Changes in cortical activity during mental rotation: a mapping study using functional MRI. Brain. **119** (1996) 89–100.
2. Kosslyn, SM., et al.: Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. Neuroreport. **12** (2001) 2519–2525.
3. Sack, AT., et al.: The dynamics of interhemispheric compensatory processes in mental imagery. Science. **302** (2005) 702–704.
4. Ogawa, K., Inui, T.: Lateralization of the posterior parietal cortex for internal monitoring of self-versus externally generated movements. J. Cogn. Neurosci. **19** (2007) 1827–1835.
5. Schenden, H.E., Stern, C.E.: Where vision meets memory: prefrontal–posterior networks for visual object constancy during categorization and recognition. Cereb. Cortex. **18** (2008) 1695–1711.
6. Ogawa, K., et al.: Neural correlates of state estimation in visually guided movements: an event-related fMRI study. Cortex. **43** (2007) 289–300.

An fMRI Investigation of the Mental Perspective Shift in Language Comprehension

Toshiki Iwabuchi, Masato Ohba, Toshio Inui, and Kenji Ogawa

Abstract It has been hypothesized that the subject, at the onset of a sentence, usually determines the perspective of the reader in sentence comprehension. We used functional magnetic resonance imaging (fMRI) to investigate the neural basis underlying the mental perspective shift in sentence comprehension. Participants observed an event consisting of two objects, immediately after their perspective was guided to one particular object by a spatial cue. The participants then judged whether the displayed sentence correctly described the previous event. If the subject of the sentence did not match the reader's predetermined perspective, then a mental perspective shift should occur. We found that the right dorsolateral prefrontal cortex (dlPFC) showed greater activation under the invalid-cue conditions requiring the mental perspective shift. The right dlPFC might therefore subserve the mental perspective shift in sentence comprehension.

Keywords Sentence comprehension · Mental perspective shift · functional MRI

1 Introduction

It has been hypothesized that the subject, at the onset of a sentence, usually determines the perspective of the reader or listener in sentence comprehension [1]. Furthermore, from a perspective of Cognitive Linguistics, a subject is a nominal element that elaborates the *trajector* of the relationship profiled, which is characterized as the initial focus of attention in a chain representing the order in which elements are accessed [2]. According to these ideas, when the subject of the presented sentence does not match the reader's or the listener's predetermined perspective, a kind of mental perspective shift should occur.

T. Iwabuchi (✉)
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
e-mail: iwabuchi@cog.ist.i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_56,

This mental perspective shift requires a greater processing cost in sentence comprehension [3]. The present study investigated the neural basis underlying the mental perspective shift using fMRI. Our experiment consisted of two tasks. In Task 1, we tried to measure the brain activity when a presented subject differed from the reader's predetermined perspective. Task 2 was designed to exclude the possibility that the brain activity recorded in Task 1 reflected eye movements or attention switching.

2 Materials and Methods

2.1 Participants

Seventeen healthy native Japanese-speaking subjects participated in this study (9 male, mean age 26 years, right-handed). Due to technical problems, four of 17 participants had to be excluded from the analysis.

2.2 Task Procedures

Our experiment consisted of two sessions. In both sessions, a spatial cue, an event picture, a visual masking picture, and a simple Japanese sentence were sequentially presented on the screen (Fig. 1). First, a spatial cue (crosshair) was displayed to guide the participants' perspective to one particular object. Then, the event was presented, which was a movie where a red and a blue object moved to the left or the right together on the screen. These objects were a circle or a triangle. Finally, a Japanese sentence was presented, comprised of two nominal phrases. Each phrase concerned one particular object that had been represented as a colored circle or triangle, together with a case particle (*-ga*, *-wo* or *-ni* corresponding to "nominative", "accusative", and "dative" designations, respectively). Verbs were chosen from "*osu*" or "*hiku*" ("push" and "pull", respectively). If the subject of this sentence did not match the participants' predetermined perspective – in other words, if the cue was invalid – a kind of mental perspective shift should occur. Furthermore, we used

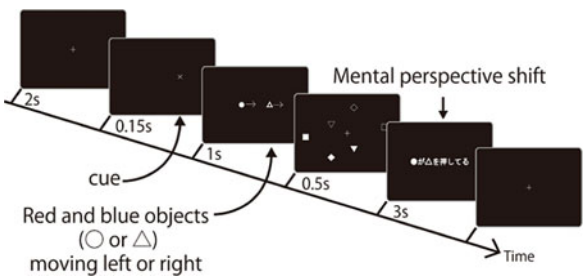


Fig. 1 Schematic illustration of the experiment. After a fixation cross disappeared, a cue, an event of moving picture, a visual masking, and a sentence describing the previous event were sequentially presented

not only active sentences (e.g., "*AKA-ga AO-wo oshiteru*: *The red circle is pushing the blue triangle*") but also passive sentences (e.g., "*AKA-ga AO-ni hikareru*: *The red circle is pulled by the blue triangle*"). Therefore, the experiment consisted of four conditions: the valid cue and the active sentence (VA), the invalid cue and the active sentence (IA), the valid cue and the passive sentence (VP), and the invalid cue and the passive sentence (IP).

Participants performed Task 1 and Task 2, respectively, in the first or second session. The order of tasks was counterbalanced across participants. In Task 1, participants were required to judge whether the displayed sentence correctly described the event shown in the previously presented event, by pressing a button during sentence presentation (3 s). In Task 2, participants were required to judge whether the subject of the sentence had appeared in the previous event during sentence presentation.

2.3 Functional MRI and Image Analysis

A 1.5-T MRI scanner was used to obtain images. Image processing and statistical analysis were performed using SPM5. Activation was thresholded at $p < 0.001$, uncorrected for multiple comparisons for the voxel level and the extent threshold was set at 20 contiguous voxels.

3 Results and Discussion

To investigate the performance in tasks (shown in Table 1), Analyses of variance (ANOVAs) on the accuracy and the reaction time (RT) were conducted. These included the within-subject factors TASK (Task 1, Task 2), CUE (valid, invalid), and VOICE (active, passive). Regarding RTs, a significant effect of VOICE and a significant interaction between TASK and VOICE were found ($F(1, 12) = 39.0$, $p < .001$, $F(1, 12) = 15.9$, $p < 0.005$). A simple main effect of VOICE within Task 1 was significant ($F(1, 24) = 52.1$, $p < .001$), indicating that passive sentences are more difficult to process than are active sentences. Furthermore, a main effect of CUE and an interaction between CUE and TASK were marginally significant ($F(1, 12) = 3.769$, $F(1, 12) = 3.470$, $p < .10$). Within Task 1, a simple main effect of CUE was significant ($F(1, 24) - 7.219$, $p < .05$). This analysis indicates

Table 1 Mean reaction time (RT) and mean percent of correct responses for each condition

	Condition	VA	IA	VP	IP
Task 1	RT (s)	1.57 (0.37)	1.61 (0.40)	1.76 (0.44)	1.84 (0.48)
	Accuracy (%)	93.7 (7.64)	90.4 (10.5)	89.4 (9.68)	88.7 (11.6)
Task 2	RT (s)	1.01 (0.23)	1.04 (0.27)	1.09 (0.30)	1.06 (0.26)

Standard deviations are given in parentheses

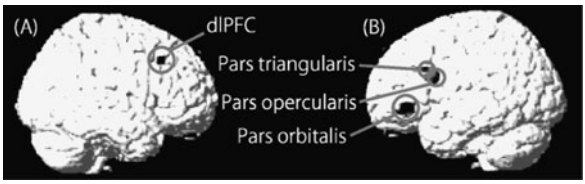


Fig. 2 **a** Greater activation in the invalid-cue conditions compared with the valid-cue conditions. **b** Greater activation in the passive conditions compared with the active conditions ($p < 0.001$ uncorrected, cluster size ≥ 20 voxels)

that the sentence comprehension cost for participants was greater when the cue was invalid than when it was valid, irrespective of voice of the sentence.

In Task 1, during sentence comprehension (for 1 s from the sentence presentation), the right dorsolateral prefrontal cortex (dlPFC) was more activated when the cue was invalid than when it was valid (Fig. 2a). Additionally, in Task 1, the pars opercularis, the pars triangularis, and the pars orbitalis of the left inferior frontal gyrus showed greater activation in the passive conditions than in the active conditions (Fig. 2b).

We then defined the activated cluster in the right dlPFC as a functional region of interest (ROI), and performed a ROI analysis. Percent signal changes within the ROI were calculated, then submitted to an ANOVA with the factors TASK, CUE and VOICE. A significant main effect of TASK and a significant interaction between CUE and TASK were found ($F(1, 12) = 11.5,\ p < .01,\ F(1, 12) = 4.87,\ p < .05$). Within Task 1, a simple main effect of CUE was significant ($F(1, 24) = 6.88,\ p < .05$). This analysis revealed that the activation of the right dlPFC was specific for Task 1.

In Japanese passive sentences, verbs are obligatorily marked with a passive morpheme *–rare*. Moreover, it has been considered that syntactic movements take place in passive sentences. Because of these processes, the passive sentences were more difficult to process than were the active sentences. The activations of the left inferior frontal gyrus observed in the present study probably reflect processing of verbal inflection [4], or processing of syntactic movements [5].

In conclusion, our study suggests that the dlPFC subserves the mental perspective shift in sentence comprehension. By comparing Task 1 with Task 2, we could exclude the possibility that the activity in the dlPFC recorded in Task 1 reflected eye movements or attention switching, which are common in the two tasks.

References

1. MacWhinney, B.: Starting points. Language. **53** (1977) 152–168.
2. Langacker, R.W.: Conceptualization, symbolization, and grammar. In Tomasello, M., ed.: The New Psychology of Language: Cognitive and Functional Approaches to Language Structure. Mahwah, NJ and London: Erlbaum (1998) 1–39.
3. MacWhinney, B.: Basic syntactic processes. In Kuczaj, S., ed.: Language Acquisition: Volume 1, Syntax and Semantics. Hillsdale, NJ: Erlbaum (1982) 73–136.

4. Yokoyama, S., Miyamoto, T., Riera, J., Kim, J., Akitsuki, Y., Iwata, K., Yoshimoto, K., Horie, K., Sato, S., Kawashima, R.: Cortical mechanisms involved in the processing of verbs: an fMRI study. J. Cogn. Neurolinguistics. **18** (2006) 1304–1313.
5. Grodzinsky, Y.: Variation in Broca's region: preliminary cross-methodological comparison. In Jenkins, L., ed.: Variation and universals in biolinguistics. Amsterdam: Elsevier (2004) 171–193.

Neural Representation of Space for Sound-Source Localization: An fMRI Study

Takafumi Sato, Toshio Inui, Hiroshi Shibata, and Kenji Ogawa

Abstract Auditory perception is important in human spatial recognition as well as in visual perception. This is especially true for the space behind the body, where visual cues are not available. This study used functional magnetic resonance imaging (fMRI) to investigate brain areas involved in processing of sound coming from different directions, including this back space. We positioned speakers at seven locations in the MRI scanner, and the subjects performed a sound localization task. Greater activation in the right superior parietal lobule (Broadmann area [BA] 7) and inferior parietal lobule (BA40) were observed when the sound was delivered from the right space compared with the left, indicating the involvement of these regions for sound localization of the right space. Activation in the right precentral gyrus (BA6) was also seen in the interaction of right/left and front/back, indicating the role of this region for sound processing of the right-front and left-back space.

Keywords Spatial cognition · Auditory localization · fMRI

1 Introduction

The location of various objects relative to one's own body need to be represented topographically in the brain in order to make smooth and successful orienting or defensive responses in a complex environment [1]. In this study, we investigated how the 3-D world around the body is represented in the brain, especially the space behind the body, where visual cues are not available. We were able to identify the brain region involved in the representation of the peripersonal space using a sound localization task [2].

T. Sato (✉)
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
e-mail: tsato@cog.ist.i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_57,

Recently, neural substrates underlying human auditory space perception has been investigated with fMRI experiments. Virtual sound sources were presented in the frontal space over headphones using sound stimuli convolved with individual head-related transfer functions (HRTFs) [3, 4]. Although sound sources can be presented virtually over headphones, the perceived object is not the sound source but a sound image. Furthermore, properties of extrapersonal space behind the back of the subject have not yet been investigated. In the present experiment, the sound stimulus was presented directly from differently-located speakers arranged around subject's head, including the space out of view.

2 Materials and Methods

A total of 22 healthy volunteers (15 males and 7 females; mean age 26 years old) participated in this study. Of these, we analyzed ten (7 males and 3 females) subjects whose accuracy rate was over 85%.

In the experiment, an auditory stimulus rotating 360° around subject's head was presented (6 s), followed by a period of silence (1 s). The starting and fading location of the rotating sound was the same, which was pseudo-randomly chosen among four possible locations (combination of left/right by front/back). A stationary sound (3 s) was then presented where the previous rotating sound had faded. Subjects were required to locate the sound by pressing buttons immediately after the sound had ended (Fig. 1). The MRI-compatible speakers were placed on the head coil of the MRI scanner at 7 orientations around subject's head except for a back position. We used white noise as an auditory stimulus with a real sound source. Each orientation was presented 12 times per session, and each subject underwent 3 sessions with the same task.

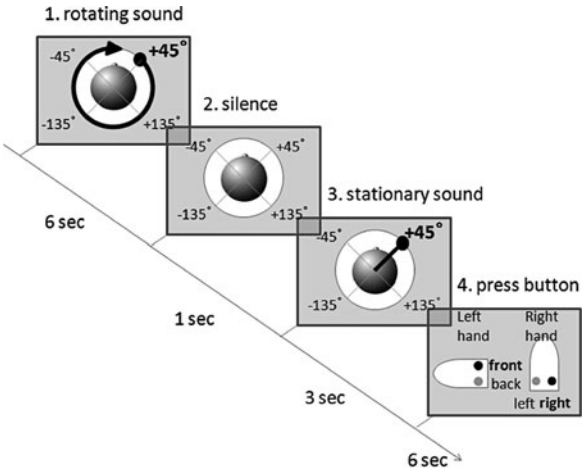


Fig. 1 The time-course with sample orientation used in the experiment

A 1.5-T scanner (Magnex, Shimadzu) was used to acquire T2*-weighted echo planar imaging with resolution of 4 × 4 × 5 mm. fMRI data analysis was performed with SPM5. We modeled the periods of stationary sound presentation (3 s) of four conditions (right-front, right-back, left-front, and left-back) as separate box-car regressors that were convolved with a canonical hemodynamic response function. Error and no-response trials were excluded from the regressors. We then compared brain activities within this period among four conditions. A second-level random-effects group analysis was conducted. Activations were thresholded at $p < 0.001$, uncorrected for multiple comparisons for the voxel level, together with an extent threshold of 20 voxels.

3 Results and Discussion

The right superior parietal lobule (BA7) and right inferior parietal lobule (BA40) were significantly more activated when the sound was presented from right directions than from left directions (the main effect of right-left direction) (Fig. 2a). These brain regions are related to the processing of right space. In addition, no significant activation was observed during the stimulus presentation from left directions compared with right. The right space predominance in human spatial recognition is consistent with the results of a previous study [6]. We also tested the interaction of right-left direction × front-back direction ([right-front > right-back] > [left-front > left-back]) and found a significant increase in the signal in the right precentral gyrus (BA6) (Fig. 2b). We then calculated percentage signal changes, defining the whole group of activated clusters in this region as a region of interest (ROI). Two-way repeated measures analysis of variance (ANOVA) of the percentage signal changes indicated a significant interaction ($p < 0.001$) (Fig. 3). The simple main effect of the right–left direction was significant in the front direction ($p < 0.001$) and was marginally significant in the back direction ($p < 0.1$). The simple main effect of the front-back direction was marginally significant in the left direction ($p < 0.06$). These results suggest that this region is related to the spatial processing of right-front and left-back spaces. This is congruent with the previous study, which indicates that the right precentral gyrus merges oculomotor and somatomotor space coding [6].

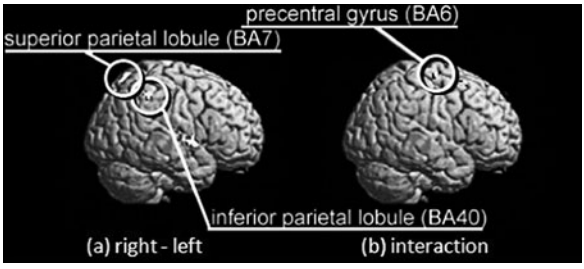


Fig. 2 Brain activity observed in the current study

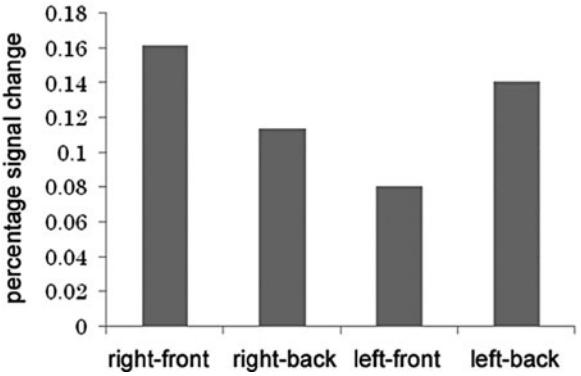


Fig. 3 The percentage signal changes in the right precentral gyrus (BA6)

In conclusion, we investigated the recognition of the space close to the head, including the back space, using the auditory space localization task and showed activation in the brain regions related to spatial processing. In particular, we found greater activation in the right superior parietal lobule and inferior parietal lobule during the sound presentation of right space, and in the right precentral gyrus during the sound presentation of right-front and left-back spaces. These regions therefore appear to play essential roles in spatial recognition.

Acknowledgment This research was supported in part by a grant from the Genesis Research Institute.

References

1. Graziano, M.S.A., Cooke, D.F., Taylor, C.S.R.: Coding the location of the arm by sight. Science. **290** (2000) 1782–1786.
2. Previc, F.H.: The neuropsychology of 3-D space. Psychol. Bull. **124**(2) (1998) 123–164.
3. Brunetti, M., Belardinelli, P., Caulo, M., Gratta, C.D., Penna, D., Ferretti, A., Lucci, G., Moretti, A., Pizzella, V., Tartaro, A., Torquati, K., Olivetti Belardinelli, M., Romani, G.L.: Human brain activation during passive listening to sounds from different locations: an fMRI and MEG study. Hum. Brain Mapp. **26** (2005) 251–261.
4. Lewald, J., Riederer, K.A.J., Lentz, T., Meister, I.G.: Processing of sound location in human cortex. Eur. J. Neurosci. **27** (2008) 1261–1270.
5. Misaki, M., Matsumoto, E., Miyauchi, S.: Asymmetric activation to the context-dependent right in the right inferior frontal region. Eur. J. Neurosci. **19** (2004) 1425–1429.
6. Iacoboni, M., Woods, R.P., Lenzi, G.L., Mazziotta, J.C.: Merging of oculomotor and somatomotor space coding in the human right precentral gyrus. Brain. **120** (1997) 1635–1645.

Investigating Causality Between Interacting Brain Areas with Multivariate Autoregressive Models of MEG Sensor Data

George Michalareas, Jan Mathijs Schoffelen, and Joachim Gross

Abstract In this work we investigate the feasibility of building a MAR model directly on MEG sensor measurements and projecting the model in brain locations where causality is calculated through Partial Directed Coherence (PDC). This method overcomes the problems of model non-robustness and large computation times encountered during causality analysis by existing methods, which first project entire MEG sensor time-series into a large number of brain locations and then the MAR model is built on this large number of time-series.

1 Introduction

In recent years there has been significant effort in developing methods for identifying causality in information flow between brain areas. A family of such methods study the frequency characteristics of Multivariate Autoregressive Models (MAR) built on time-series of activated brain sources [1], with the most widely used being Partial Directed Coherence [2]. These time-series are reconstructed from MEG sensor data through spatial filters derived commonly by beamforming algorithms [3]. The large number of potential activation sources produced by such algorithms corresponds to a large number of activation time-series, which is prohibitive for the derivation of robust MAR models. In this work we investigate the derivation of MAR model directly on MEG sensor data and its projection in the source space. By this method the modelling process is performed on the sensor space which has moderate dimensionality as compared to the high-dimensional source space. This leads to greater model robustness as well as significantly reduced computation times. Feasibility of a similar approach for EEG data has already been shown by Gómez-Herrero et al. [4].

G. Michalareas (✉)
Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, UK
e-mail: georgem@psy.gla.ac.uk

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_58,

2 Multivariate Modelling of MEG Data

If $\psi(t)$ are the activation signals in brain space and $\mathbf{x}(t)$ the sensor measurements in the sensor space corresponding to the activation signals then the following relationship holds:

$$\mathbf{x}(t) = \boldsymbol{\Lambda}\psi(t) + \boldsymbol{\nu}(t) \tag{1}$$

where $\boldsymbol{\Lambda}$: Leadfield Matrix (Depends on brain geometry and volume conductor model) and $\boldsymbol{\nu}$: Measurement noise at sensor space, assumed to be White Gaussian. Biological noise representing processes not related to brain activity (i.e. fields from muscle activity) is assumed negligible (already filtered out) here and is thus omitted. For an arbitrarily large number of potential activation sources and their corresponding time series described by ψ, a multivariate model can be described as:

$$\psi(t) = \sum_{r=1}^{p} \mathbf{B}_{\psi}(\tau)\psi(t-\tau) + \boldsymbol{\epsilon}(t) \tag{2}$$

where p: model order, τ: delay step, $\mathbf{B}_{\psi}(\tau)$: model coefficients matrix for delay step τ and $\boldsymbol{\epsilon}$:multivariate residual. Combining (1) and (2) gives the following expression for the multivariate model in the Sensor Space:

$$\mathbf{x}(t) = \sum_{r=1}^{p} \boldsymbol{\Lambda}\mathbf{B}_{\psi}(\tau)\boldsymbol{\Lambda}^{+}\mathbf{x}(t-\tau) + \boldsymbol{\Lambda}\boldsymbol{\epsilon}(t) \tag{3}$$

where $^{+}$ denotes the Moore-Penrose pseudoinverse. By applying Principal Component Analysis (PCA), collinearity can be eliminated and insignificant Principal Components can be omitted. Then the model in eqref(modeleq3) becomes:

$$\mathbf{x}_{PCA}(t) = \sum_{r=1}^{p} \boldsymbol{\Phi}\boldsymbol{\Lambda}\mathbf{B}_{\psi}(\tau)(\boldsymbol{\Phi}\boldsymbol{\Lambda})^{-1}\mathbf{x}_{PCA}(t-\tau) + \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\epsilon}(t) \tag{4}$$

where $\boldsymbol{\Phi}$: Matrix of Significant Feature Vectors for Principal Components.

If a multivariate model is developed directly on the Principal Components of the MEG sensor measured time-series,it is evident from (4) that if the Leadfield Matrix $\boldsymbol{\Lambda}$ and the Feature Vectors Matrix $\boldsymbol{\Phi}$ are known, then the coefficient matrix $\mathbf{B}_{\psi}$ and the residual $\boldsymbol{\epsilon}(t)$ in the brain source space can be estimated.

The Feature Vectors Matrix is based on the selection of the eigenvectors corresponding to the highest eigenvalues. Commonly the eigenvectors selected can explain 95% of the variance. The Leadfield Matrix $\boldsymbol{\Lambda}$ can be estimated in two different ways. Beamforming algorithms applied on the MEG sensor measurements and a predefined grid of voxels in the brain provide an estimation of the dipole orientation vector for each voxel and a Spatial Filter, mapping a measurement of each MEG

sensor to the 3 orthogonal components of the activation signal in each voxel. For the same predefined voxel grid, the reverse map from the 3 orthogonal components of the activation signal to each MEG sensor measurement can be precomputed by running the forward problem. The inverse of the Beamformer's Spatial Filter should approximate this 3-dimentional Leadfield for each MEG sensor. The 1-dimentional Leadfield Matrix $\boldsymbol{\Lambda}$ for each voxel can then be estimated in two ways, by combining the estimated orientation vector with either the 3-dimentional precomputed Leadfield or with the inverse of the Spatial Filter.

After choosing the Feature Vector Matrix and computing the Leadfield Matrix, the coefficient matrix $\mathbf{B}_{\psi}$ and the residual $\boldsymbol{\epsilon}(t)$ in the brain source space are estimated. Directionality analysis is based on analysing the coefficients of the estimated $\mathbf{B}_{\psi}$ in the frequency domain. The most widely used method for such analysis is Partial Directed Coherence.

Partial Directed Coherence is a method used in multivariate systems to identify directionality in Coherence between pairs of variables. It operates on the coefficients of multivariate models derived on corresponding multivariate data. For a given Voxel Pair *ij* and a given frequency *f* PDC is effectively the FFT of the series of model coefficients forecasting *i* from *j* for various time lags, normalized by the norm of the FFT of the model coefficients for all pairs with *j* as the causal pair. PDC thus in normalized with respect to the transmitting voxel j and takes values between 0 and 1, the latter describing high causality.

3 Simulated Brain Dipole

MEG sensor data was simulated. Five activation signals inside the brain volume were generated from MAR equations approximating damped oscillators [1]. Activation signals have predefined causality (Fig. 1). The overall Brain dipole field was simulated through Fieldtrip Toolbox in MATLAB and White Gaussian Noise representing noisy environment processes was added. The activation signals were designed with nominal frequency 10 Hz (Alpha band) and sampling frequency 100 Hz. Time-Series were organized in 20 Trials with each trial generated from a random impulse. Figure 2 shows the impulse response of the signal generators.

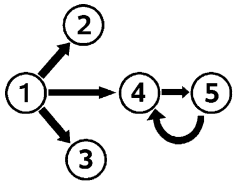


Fig. 1 Predefined causality of simulated activation signals

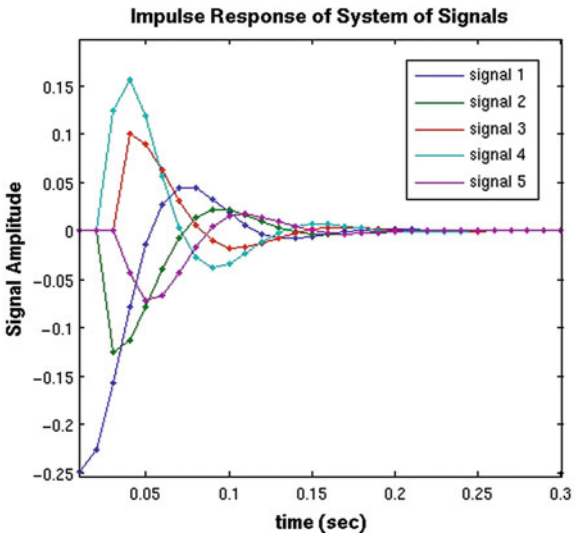


Fig. 2 Impulse response of the system of activation signals

4 Confidence Intervals of PDC

As PDC depends on the spectrum of the estimated MAR model coefficients for each pair, it is necessary to be accompanied by confidence intervals. Three different methods were used.

- Shuffling all data points [5]: All data points in the time-series were randomly shuffled. This process destroys completely causal information. The shuffling process was repeated 200 times. This process in effect provides the range of PDC for a random signal with the same distribution as the analyzed time-series.
- Shuffling Trials [5]: Data within each Trial remains fixed. Trials are randomly shuffled separately for each time-series. The process was repeated 200 times. This process in effect examines the boundaries of PDC when phase between time-series is randomized while causality structure is preserved.
- Jackknife across Trials [6]: One trial was removed from all time-series (the same across time-series) and PDC was estimated. The process was repeated 20 times (the number of trials). This process in effect examines if causality is dominantly affected by a single trial.

5 Stationarity Assessment

In order to investigate the level of determinism in the MEG data (simulated and real) to be used for the MAR models, we assesed its stationarity through the criterion of second-order weak stationarity. As MEG data is primarily obtained in trials with each trial typically including a single occurence of the process under investigation,

the epoch bin selection criterion was applied on a per trial basis. The mean and variance were calculated for each trial and for the entire time-series. Next, we compared the statistical parameters of the total time series with those of each segment. If these differences were significant with probability ¡0.05.the segment would be discarded but this did not happen in any of the 20 trials unlike what normally happens with real data. Intra-trial stationarity has not been addressed in this work as the linear system of dumped oscillators had a definite deterministic nature. The sampling frequency of the data is 100 Hz and such low sampling rates could cause uncertainty regarding stationiarity assessment, especially when small time intervals are recording per trial. For intra-trial stationarity assessment, the best approach is to acquire data with a high sampling rate, typically 2 Khz for MEG nowdays.

6 Results

6.1 PDC from Projected MAR Model Is Close to Ideal

Ideal reference PDC was calculated directly from the 5 simulated activation signals. It is shown in Fig. 3a for all possible pair combinations between signals. The pairs that show distinct PDC are in agreement with the expected causal pairs from the configuration of the simulated activation signals. These pairs are: 21, 31, 41, 54, 45. Then the PDC was calculated from MEG sensor data. After the MAR model was built on the Principal Components of MEG Sensors Time-series, it was projected in 5 locations close to the points of the 5 simulated activation signals with (i): Inverse Leadfield and (ii): Spatial Filter estimated by a Beamformer. PDC calculated for case (i) is shown in Fig. 3b and for case (ii) in Fig. 3c. PDC calculated from the Projected MAR Model in approximating the Ideal reference PDC from the activation signals. This means that a MAR model built from MEG data in the Sensor Space and projected in the Source Space preserves causality information for the underlying generating activation processes. Examination of the levels of PDC shows that it is distinguishably high for the Causal Pairs compared to all the rest non-causal pairs. Maxima occur around 10 Hz, the nominal frequency of the activation signals.

6.2 PDC Deviation from Ideal w.r.t Environment Noise, N of Samples and Model Order

We investigated the deviation of PDC from the "ideal" for different values of the following parameters:

– Environment Noise: White Gaussian Noise added to simulated Dipole. Amplitude is relative to Dipole Amplitude. Investigated Range: 1–20. This number is the rate of power of the noise dipole, superimposed on the dipole generated from the simulated signals and measured and the locations of the MEG sensors. The above

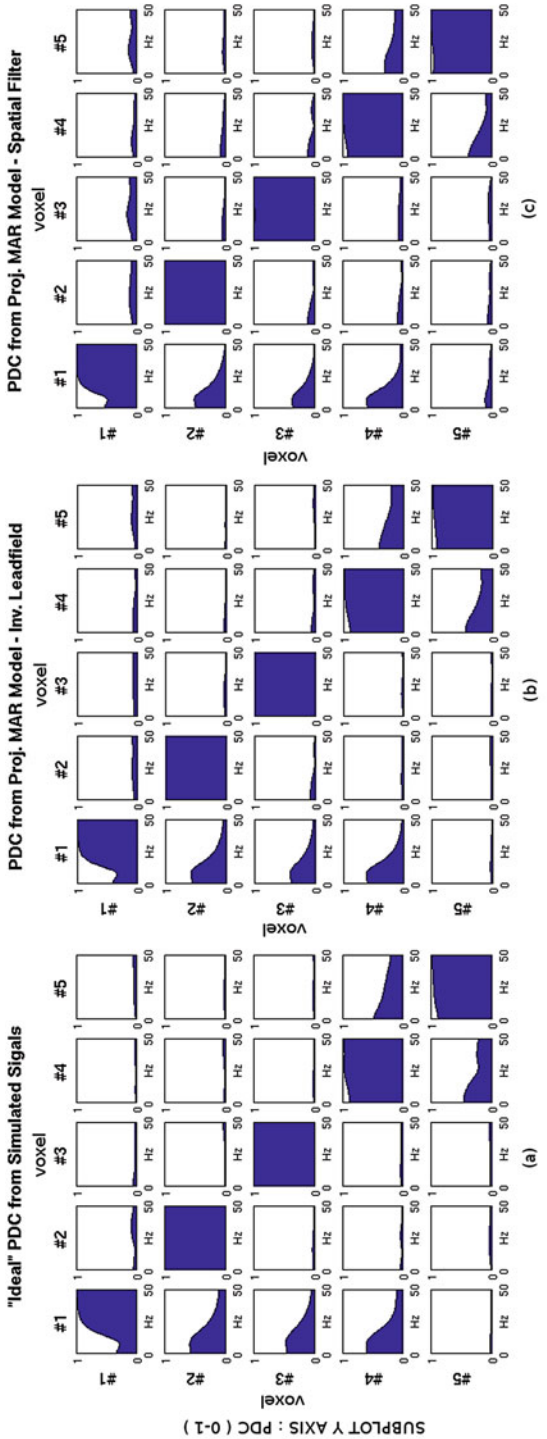


Fig. 3 **a** "Ideal" PDC directly from activation signals **b** PDC from MAR model projected by Inverse leadfield **c** PDC from MAR Model projected by spatial filter

type of noise was chosen because it was attempted to simulate real MEG lab noise which usually follows such a distribution due to the superposition of multiple individual noise sources. Testing of different types of noise could provide a good insight in investigation of mainly single noise source senarios either of biological or artificial origin.
– N Samples/Trial: 20 Trials were used. Investigated range: 500–4,000
– MAR Model Order: Number of time-lags used in the model. Investigated Range: 5–100

The objective here was to investigate average PDC deviation from Ideal for the above parameters, around the frequency of the activation signals (10 Hz) and separately for Causal and Non-Causal Pairs. For this purpose two metrics were used

– Metric 1: mean(PDC of Proj. MAR Model Ideal PDC) for freq. 7–12 Hz averaged for the Causal Pairs 21, 31, 41, 54, 45.
– Metric 2: same as Metric 1 but for the rest, Non-Causal Pairs.

In Fig. 4a the above 2 metrics are shown in a comprehensive way describing the variation of all three investigated parameters. Each subplot presents the 2 metrics vs N Samples/Trial. Each suplot corresponds to a different combination of Environment Noise Level and MAR Model Order. Up to Noise Levels 4 times the strength of the dipole the deviation of the PDC for the Causal Pairs remains low. In the presence of higher noise, PDC is deviating significantly from the ideal. When a small N of Samples is used and high Model Order, PDC for both Causal and Non-Causal pairs is inconsistent relative to the ideal. In such cases one cannot distinguish between Causal and Non-causal pairs. Each subplot presents the 2 metrics vs N Samples/Trial. Each suplot corresponds to a different combination of Environment Noise Level and MAR Model Order. Up to Noise Levels 4 times the strength of the dipole the deviation of the PDC for the Causal Pairs remains low. In the presence of higher noise, PDC is deviating significantly from the ideal. When a small N of Samples is used and high Model Order, PDC for both Causal and Non-Causal pairs is inconsistent relative to the ideal. In such cases one cannot distinguish between Causal and Non-causal pairs.

6.3 More Principal Components Do Not Improve PDC

PCA was applied to the 248 MEG sensor time-series. The first 5 Principal Components explained 95% of the variance. We investigated the average deviation of PDC from the Ideal for Causal and Non-Causal Pairs with respect to the number of Principal Components used in the MAR Model. Investigated Range: 5–248 Principal Components. Figure 4b presents the results for Causal (green) and Non-Causal Pairs(red). PDC was not improved by the inclusion of more Principal Components.

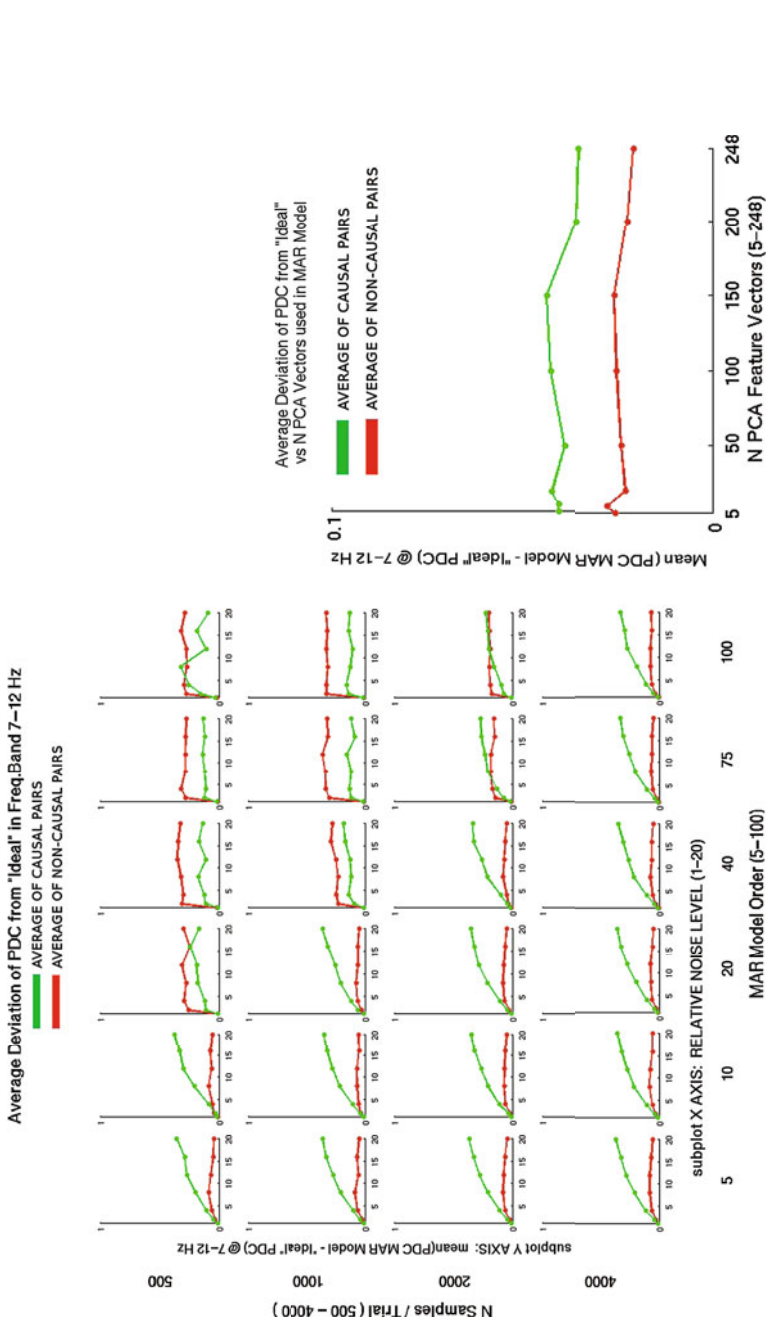


Fig. 4 **a**. Average deviation of PDC from "Ideal" for causal and non-causal pairs. In subplot PDC is plotted vs N samples/trial. Each subplot corresponds to a different combination of environment noise level and MAR model order. **b** Average deviation of PDC from "Ideal" for causal and non-causal pairs vs number of principal components used in the MAR model

6.4 Confidence Intervals for PDC

Shuffling all data points provides the layer of random PDC close to 0. Shuffling Trials provides the PDC distorted by Phase Noise. In these 2 cases significant PDC should be outside these boundaries. Jackknife provides a layer around PDC which should not include 0 if PDC is to be significant. As it can be seen from Fig. 5 the

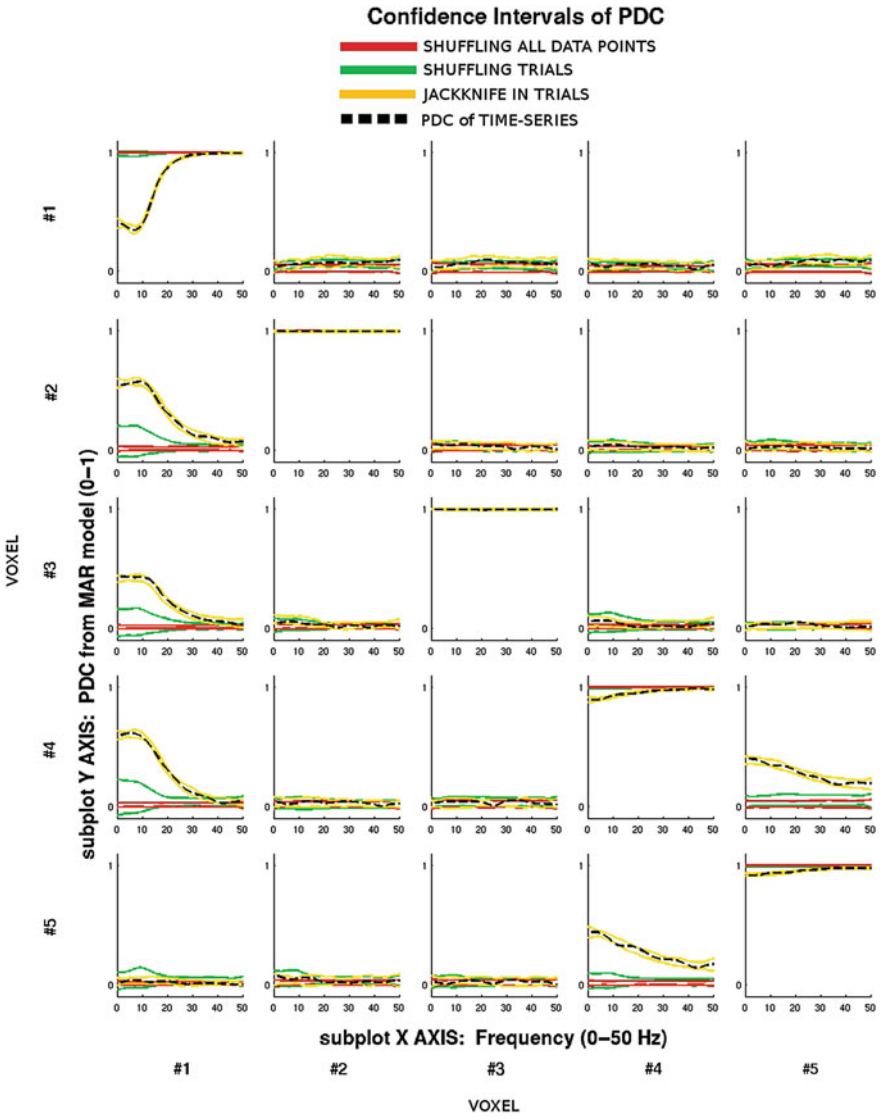


Fig. 5 Confidence intervals of PDC by 3 methods: shuffling all data points, by shuffling trials, by Jackknife across trials

PDC for the causal pairs can be clearly distinguished when confidence intervals are examined. For the non-causal pairs, the PDC and the confidence intervals overlap and in these cases PDC can be characterized as “insignificant”.

7 Conclusions

Building a MAR model directly on MEG sensor data and projecting it in Brain locations is, in terms of causality content, identical to building the model in Brain space from projected time-series. Causality, as measured by PDC, estimated from projected MAR model is approximating the Ideal, evaluated directly from simulated activation signals in the Brain. This fact created the advantage of building the MAR Model directly on the MEG sensor time-series and then have the flexibility to project only the MAR model coefficients in an arbitrary number of locations inside the brain. Due to the low dimensionality of the MAR Model coefficient series as compared to the entire time-series, the computational load for this projection is far smaller than existing methods which first project the MEG sensor time-series and then estimate the MAR Model.

Estimation of Causality through PDC, is tolerant of Environment Noise, but suffers from small number of samples is combined with large Model Order. Using only the significant Principal Components of MEG sensor data in the MAR model, is adequate. Inclusion of more Principal Components does not improve PDC. Confidence Intervals calculated by Shuffling All Data Points, Shuffling Trials and Jackknife on Trials show that Causal Pairs and Non-Causal Pairs can be confidently distinguished.

References

1. Baccalá, L.A., Sameshima, K.: Partial Directed Coherence: a new concept in neural structure determination. Biol. Cybern. **84** (2001) 435–439.
2. Gómez-Herrero, G., Atienza, M., Egiazarian, K., Cantero L.C.: Measuring directional coupling between EEG sources. Neuroimage (2008) doi:10.1016/j.neuroimage.2008.07.032.
3. Gourévitch, B., Bouquin-Jeannes, R.L., Faucon, G.: Linear and Nonlinear Causality Between Signals: Methods, Examples and Neurophysiological Applications Comparative Genomics. Dordrecht, NL: Kluwer Academic Press. (2006) 47–58.
4. Hui, H.B., Leahy, R.M.: Linearly constrained beamformers for MAR modelling of cortical interactions. Biomedical Imaging:Nano to Macro, 2006, 3rd IEEE International Symposium on (2006), pp.237-240, doi: 10.1109/ISBI.2006.1624896.
5. Kaminski, M., Ding, M.Z., Truccolo, W.A., Bressler, S.L.: Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol. Cybern. **85**(2) (2001) 145–157.
6. Schlögl, A., Supp, G.: Analyzing event-related EEG data with multivariate autoregressive parameters. In Neuper, K., eds.: Progress in Brain Reseach (159). Amsterdam: Elsevier (2006) 135–147.

EEG Feature Extraction During Mental Fatigue and Relaxation by Principal Component Analysis

Lan-Lan Chen, Jun-Zhong Zou, Jian Zhang, Chun-Mei Wang, and Min Wang

Abstract EEG is one of the most predictive and reliable measurements for mental fatigue and relaxation evaluation. The aim of this study is to transform a number of EEG spectrum variables into few principal components using PCA method. After transformation, EEG multivariate dataset can be visualized in a lower-dimensional space where different mental states are clearly discriminated from each other.

Keywords Mental fatigue · Relaxation · EEG · feature extraction · PCA

1 Introduction

Mental fatigue is a curial issue of modern everyday life [1]. EEG variables can be viewed as reliable tools for mental evaluation. It is important to transform a number of possibly correlated variables into a smaller number of uncorrelated variables to provide comprehensive evaluation. PCA method supplies us a lower-dimensional description of a multivariate dataset from the most informative viewpoint [2, 3].

The flow chart for this study is illustrated in Fig. 1. EEG data was recorded at two mental states: mental fatigue caused from two 2-digit integer mental addition and mental relaxation in a classical music environment. EEG time series was transformed into power spectrum by fast Fourier transform (FFT). PCA method was applied to obtain principal components from a variety of EEG spectrum variables. After then, support machine vector was utilized to classify two mental states.

L.-L. Chen (✉)
Department of Automation, East China University of Science and Technology, Shanghai 200237, China
e-mail: chenlanlan104@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_59,

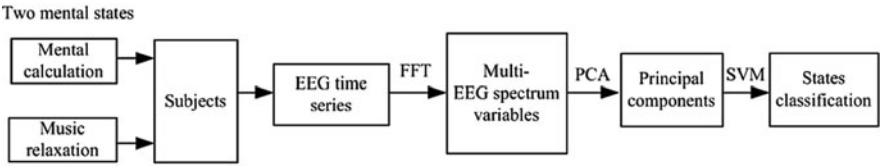


Fig. 1 The flowchart of the proposed method

2 Methods

2.1 Subjects and Recording Condition

Seven male participants, aged 22–25 years old ($M = 23$) were recruited from East China University of Science and Technology. EEG signals were recorded using Ag/AgCl electrodes. The electrical impedance was under 10 kOhms for all electrodes. EEG electrodes were placed according to the international 10–20 system and put at the following areas: Fp1, Fp2, F3, F4, Fz, Cz, O1, and O2 against ipsilateral earlobe electrode (A1, A2 or Aav). EEG signals were amplified by a digital EEG machine (Nihon-Koden EEG2110) with the sampling frequency of 200 Hz, a low-frequency cut-off of 0.5 Hz and a high-frequency cut-off of 30 Hz.

2.2 EEG Spectrum Variables

EEG data were stored by segments. Each segment contained 1,024 samples (5.12 s). The digital data of EEG time segments are transformed into the Fourier components by the fast Fourier transform (FFT) algorithm. The periodogram of each segment was obtained as the squares of the magnitude of the Fourier components divided by the data length for each resolving frequency of 0.2 Hz. EEG variables can be expressed by the periodogram components as in (1) and (2).

$$\text{Amplitude } A_z(c) = 4\sqrt{S_z(c)}, \tag{1}$$

$$\text{Percentage } P_z(c) = \frac{S_z(c)}{ST(c)} \times 100. \tag{2}$$

Where c represented each channel; z denoted the respective EEG frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–25 Hz) and total band T (0.5–25 Hz). S_z (c) was the summation of periodogram components within the frequency band of z at channel c. S_T (c) was the summation of periodogram components within the frequency band of T at channel c.

Nine variables highly related to mental fatigue, drowsy and concentration were selected for the next principal component analysis. They were:

Delta components at Fz: A_δ (Fz) and P_δ (Fz);
Theta components at Fz: A_θ (Fz) and P_θ (Fz);
Alpha components at O1: A_α (O1) and P_α (O1);
Beta components at O1: A_β (O1) and P_β (O1);
Total periodogram components at O1: A_T (O1).

2.3 Principal Component Analysis

PCA is eigenvector-based multivariate analysis. The goal is to transform a given data set X of dimension p to an alternative data set Y of smaller dimension q. Original data X is denoted by the eigen vector α_i ($i = 1, 2, \cdots, p$) and the eigen value λ as in Eq. (3).

$$X\alpha_i = \lambda\alpha_i, \tag{3}$$

The accumulative contribution rate C is calculated by all eigen vector λ as in Eq. (4).

$$C = \frac{\sum_{i=1}^{q} \lambda_i}{\sum_{i=1}^{p} \lambda_i} \tag{4}$$

Usually, q should ensure C is over 70–90% [3].

Subsequently, classification methods such as support vector machine (SVM) can be used to discriminate data set in the new q-dimension space.

3 Results

In this study, original data set X included nine EEG spectrum variables. Using PCA, X was transformed into a new coordinate system by projection of the data comes to lie on the principal components. The accumulative contribution rate of the first three principal components was 85%. Therefore, original dataset X was projected into the first three dimensions. The values in the remaining dimensions were dropped with minimal loss of information.

The data distribution in the first three principal components is illustrated in Fig. 2. Data were recorded from seven subjects under two mental states.

Training sets included 39 EEG segments, 21 segments of which were recorded at mental calculation and the left 18 segments were recorded at music relaxation.

Test sets included 26 EEG segments, 17 segments of which were recorded at mental calculation and the left 9 segments were recorded at music relaxation.

From Fig. 2, training sets and test sets were visualized in the three-dimension space. And two mental states can be clearly discriminated from one another. Using support vector machine (SVM), test sets were correctly classified.

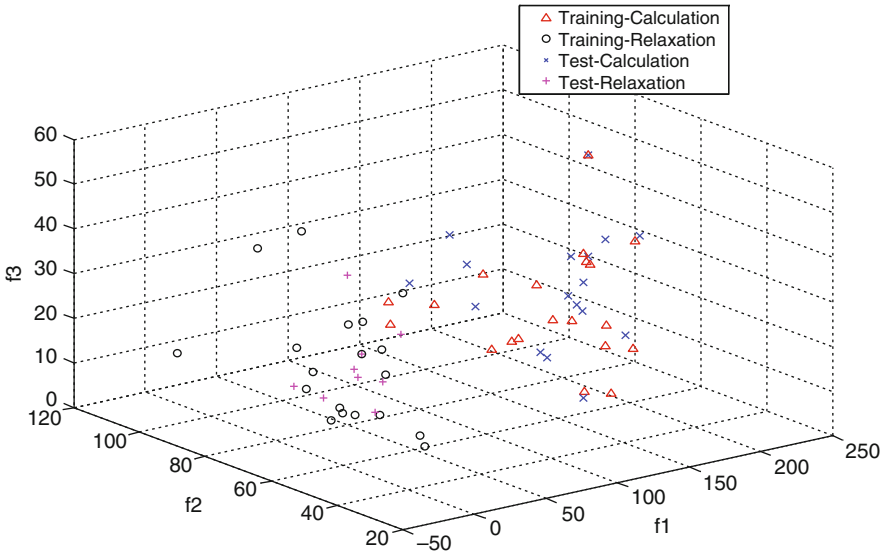


Fig. 2 The data distribution in the first three principal components. *f*1, *f*2, and *f*3 represent the coordinate of the first three principal components, respectively

4 Conclusions

EEG spectrum variables can be viewed as reliable estimators for mental fatigue and relaxation states. Principal components analysis helps to transform multivariable data set into a smaller number of principal components. After transformation, EEG can be visualized in a lower-dimension space where different mental states are easy to discriminate by classification methods such as support vector machine (SVM). In the future work, the methods proposed in this paper can be extended to visualization and discrimination of even more mental states.

References

1. Levi, L.: Guidance on Work-Related Stress: Spice of Life, or Kiss of Death? Luxembourg: Office for Official Publications of the European Communities (2000).
2. Makeig, S., Jung, T.: Changes in alertness are a principal component of variance in the EEG spectrum. Neuroreport. **7** (1995) 213–216.
3. Jolliffe, L.T.: Principal Component Analysis. New York, NY: Springer(1986).

Kernel Granger Causality Mapping Effective Connectivity: A Resting fMRI Study

Wei Liao, Daniele Marinazzo, Zhengyong Pan, and Huafu Chen

Abstract The human brain is a complex, dynamic, nonlinear system and operates far from equilibrium. The functionally connected of human brain being potentially dynamic and directional, may not be adequately captured by simple correlation, or anti-correlation. To evaluate the possible effective connectivity as well as the nonlinear effective connectivity within human brain, we applied kernel Granger causality (KGC) to the 90 cortical and subcortical regions of interest (ROIs) of resting state fMRI data. Our analysis also found the hub node that was characterized by much number of Granger causal efferent connections to this given node from any other node at different level of nonlinearity. Overall, our results revealed the causal influences and hubs among these ROI at different order of nonlinearity.

Keywords Kernel Granger causality · Resting state · Effective connectivity

1 Introduction

The human brain is a complex, dynamic, nonlinear system and operates far from equilibrium [1]. The functionally connected of human brain be usually captured by simple correlation, or anti-correlation particularly in resting state functional magnetic resonance imaging (fMRI) data [2, 3]. However, the functionally connected of human brain being potentially dynamic and directional that refers to the effective connectivity, which refers explicitly to the influence that one neuronal system exerts over another, either at a synaptic or population level, either directly or indirectly [4]. Granger causality analysis can provide information about the dynamics and directionality of fMRI signal in cortical circuits [5–7]. Alternatively, Granger causality analysis is likely to be an appropriate approach to study the directional interactions of these resting state networks [8]. This Granger causality analysis that that based on

H. Chen (✉)
Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China
e-mail: chenhf@uestc.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_60,

Geweke's dependence measure [9] indeed only considered the linear influence and hitherto ignored the nonlinear connectivity. In the present study, aimed to investigate the nonlinear connectivity of the brain, we use kernel Granger causality (KGC) method [10, 11] to explore effective connectivity during resting state fMRI data.

2 Methods

2.1 Subjects and Data Acquisition

One right-handed subject participated in this study. The present study was approved by the local ethical committee in West China Hospital of Sichuan University, and informed written consent was obtained prior to the experiment. For the resting state scans, subjects were instructed simply to rest with their eyes closed, not to think of anything in particular, and not to fall asleep. Experiments were performed on a 3.0-T GE-Signa MRI scanner. Functional images were acquired using EPI sequence (TR = 1000 ms, TE = 30 ms and flip angle = 90°). Sixteen transverse slices (FOV = 24 cm, in-plane matrix = 64 × 64, slice thickness = 6 mm, without gap) were acquired.

2.2 Data Processing

Data preprocessing was partly carried out using the Statistical Parametric Mapping software. The 200 volumes were first corrected for the temporal difference in acquisition among different slices, and then the images were realigned to the first volume for head-motion correction. The fMRI images were realigned with the corresponding T1-volume and warped into a standard stereotaxic space at a resolution of $3 \times 3 \times 3$ mm^3, using the Montreal Neurological Institute (MNI) echo-planar imaging template. Then, they were spatially smoothed (FWHM = 8 mm). For details step see Fig. 1. The more details of the model can be found in our previous studies [10, 11], and the application in fMRI data was described in our previous study [12].

3 Results

Considering KGC can control the order of nonlinearity of the fMRI time series, the linear Granger influence network (Fig. 2a) and the nonlinear Granger influence network (Fig. 2b) were gain using by KGC. After creating the brain network using the given threshold, we then determined which nodes were connected to the largest number of other nodes (using in-degree), i.e. which nodes are "hubs". Specifically, we define a hub as a node whose in-degree is larger than the average in-degree of the network (see Fig. 2 node size). In linear Granger influence network (Fig. 2a), many hub located in the frontal lobe, while in nonlinear Granger influence network

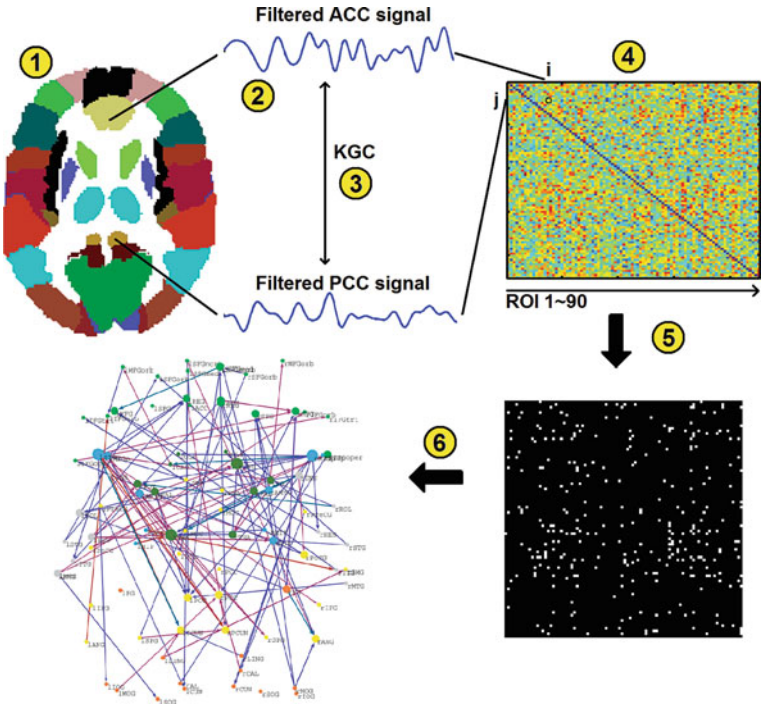


Fig. 1 Schematic illustration of the KGC analysis on resting state fMRI data. The first step (*panel 1*) the images were segmented into 90 anatomical ROI. The averaged time series of each ROI had been regression against six head motion parameters, signal in cerebrospinal fluid and in the white matter and had been temporal band-pass filtered ($0.01 < f < 0.008$ Hz) (*panel 2*). Granger influence values of each pairwise ROI were calculated using by KGC (*panel 3*). Granger influence values were entered an asymmetric directed matrix (*panel 4*). An asymmetric binary directed matrix gain after using thresholded along with FDR correction (*panel 5*). The regional centroid of each ROI (*node*) and was the arcs (effective connectivity) laid out in line with its anatomical location in MNI stereotaxic space (*panel 6*)

(Fig. 2b), the hub node not only in the frontal lobe, but also the distribution of hub node were widespread, such as, in parietal-(pre)Motor, in medial temporal lobe and subcortical regions.

4 Discussion and Conclusions

The functionally connected of human brain be usually captured by simple correlation, or anti-correlation particularly in resting state. Granger causality analysis is likely to be an appropriate approach to study the directional interactions of these resting state networks [8]. In the current study, aimed to investigate the nonlinear connectivity of the brain in the resting state, we use KGC method [10, 11] to explore effective connectivity. Our analysis provided evidence either linear or

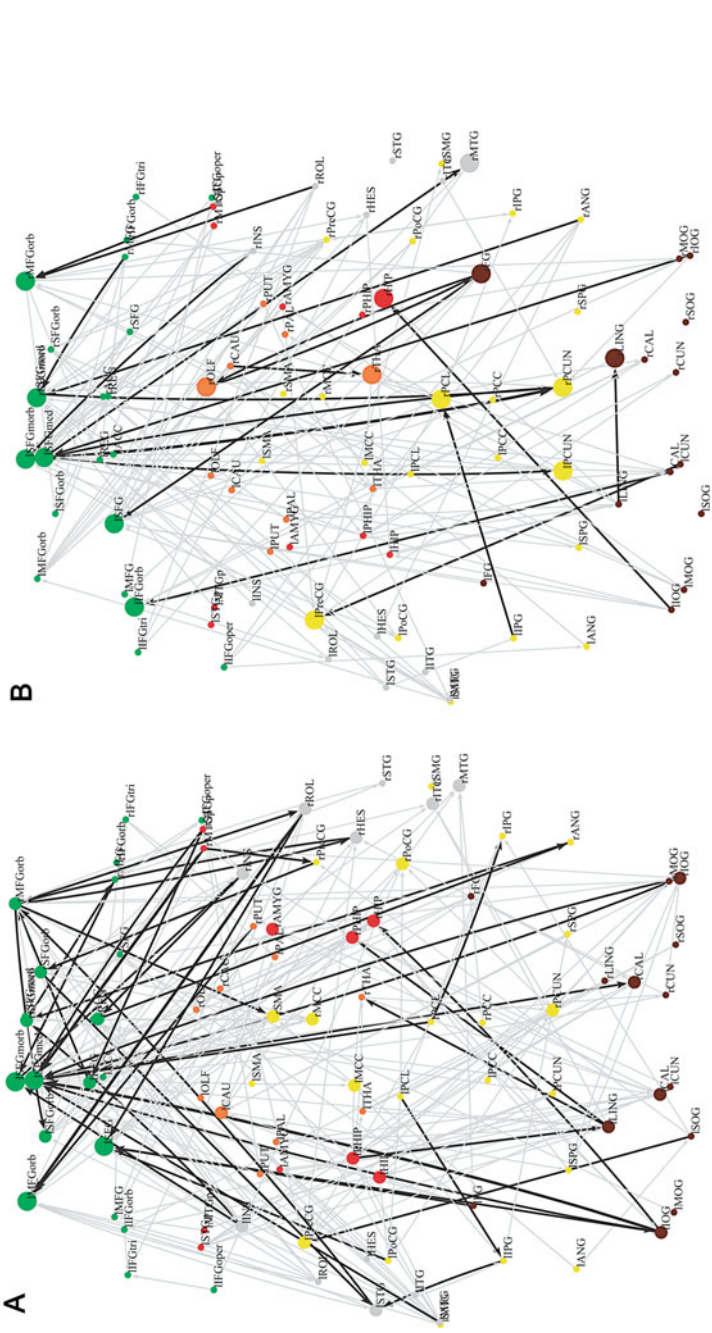


Fig. 2 Network visualization of Granger influence values matrix. The six anatomical subregions were differently colored (see legend). Node (individual ROIs) size was coded according to their in-degree. The arcs (effective connectivity) were differently colored *black* and *gray*, respectively, according to Granger influence strength ($p < 0.05$, and $p < 0.01$, FDR corrected). Granger influence network for (IP = 1) (**a**) and for (IP = 2) (**b**)

nonlinear causal influences in brain. Our analysis also found the hub node that characterize by much number of Granger causal efferent connections to this given node from any other node at different level of nonlinearity. Overall, our results revealed the causal influences and hubs among these ROI at different order of nonlinearity. Further work, however, will be needed to explore and explain the physiological significance of this effective connectivity networks.

References

1. Klonowski, W.: From conformons to human brains: an informal overview of nonlinear dynamics and its applications in biomedicine. Nonlinear Biomed. Phys. **1** (2007) 5.
2. Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. **34** (1995) 537–541.
3. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. **102** (2005) 9673–9678.
4. Friston, K.J., Frith, C.D., Frackowiak, R.S.: Time-dependent changes in effective connectivity measured with PET. Hum. Brain Mapp. **1** (1993) 69–79.
5. Gao, Q., Chen, H., Gong, Q.: Evaluation of the effective connectivity of the dominant primary motor cortex during bimanual movement using Granger causality. Neurosci. Lett. **443** (2008) 1–6.
6. Goebel, R., Roebroeck, A., Kim, D.S., Formisano, E.: Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn. Reson. Imaging **21** (2003) 1251–1261.
7. Roebroeck, A., Formisano, E., Goebel, R.: Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage. **25** (2005) 230–242.
8. Stevens, M.C., Pearlson, G.D., Calhoun, V.D.: Changes in the interaction of resting-state neural networks from adolescence to adulthood. Hum. Brain Mapp. (2009).
9. Geweke, J.: Measurement of linear dependence and feedback between multiple time series. J. Am. Stat. Assoc. **77** (1982) 304–313.
10. Marinazzo, D., Pellicoro, M., Stramaglia, S.: Kernel method for nonlinear Granger causality. Phys. Rev. Lett. **100** (2008) 144103.
11. Marinazzo, D., Pellicoro, M., Stramaglia, S.: Kernel-Granger causality and the analysis of dynamical networks. Phys. Rev. E. **77** (2008) 056215.
12. Liao, W., Marinazzo, D., Pan, Z., Gong, Q., Chen, H.: Kernel Granger causality mapping effective connectivity on fMRI data. IEEE Trans. Med. Imaging. in press.

EEG Source Localization Based on Multiple fMRI Spatial Patterns

Xu Lei and Dezhong Yao

Abstract EEG source localization is an ill-posed problem, and constraints are required to ensure the uniqueness of the solution. In this paper, using independent component analysis (ICA), multiple fMRI spatial patterns are employed as the covariance priors of the EEG source distribution. With the empirical Bayes (EB) framework, spatial patterns are automatically selected and EEG sources are estimated with Restricted Maximum Likelihood (ReML). The computer simulation suggests that, in contrast to the previous methods of EB in EEG source imaging, our approach is distinctly valuable in improvement of distributed source localization.

Keywords EEG · fMRI · Network EEG source imaging · Restricted maximum likelihood · Source reconstruction · Distributed solution

1 Introduction

Existing literatures have established a few potential protocols to employ fMRI information in EEG source localization [1]. The fMRI weighting EEG source localization [2] used the active map obtained from statistic analysis of fMRI to improve the localization accuracy. Phillips et al. [3] proposed an empirical Bayes (EB) framework to combine fMRI active map as the covariance priors controlled by hyperparameters in EEG source localization.

The statistic analysis of fMRI generally needs a proper hemodynamic response function (HRF). The difference between the actual HRF and the assumed HRF may reduce the feasibility of this method [4]. In contrast, independent component analysis (ICA) can simultaneously extract diverse spatial patterns with similar function. In addition, by relying upon an assumption of independence, ICA is intrinsically a

X. Lei (✉)
Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China
e-mail: ray_sure@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_61,

multivariate approach, and it allows the user to be agnostic with the exact form of the response.

In this paper, using spatial patterns derived from fMRI, we proposed the Network EEG Source Imaging (NESOI). With NESOI, the automatic selection of multiple spatial patterns is attainable. Comparing with other method with anatomical smoothness or sparseness constraints, NESOI can produce a solution that combines information of the high temporal resolution EEG and the high spatial resolution fMRI.

2 Methods

An EB [2, 5] model used for EEG source location is,

$$\begin{array}{ll} Y = L\theta + \xi_1 & \xi_1 \sim N(0, C_1) \\ \theta = 0 + \xi_2 & \xi_2 \sim N(0, C_2) \end{array}, \tag{1}$$

where $Y \in R^{n\times s}$ is the EEG recording with n channels and s samples. $L\square\ R^{n\times d}$ is the lead-field matrix, and $\theta \in R^{d\times s}$ is the unknown source activity for d dipoles. $N(\mu, C)$ denotes a multivariate Gaussian distribution with mean μ and covariance C. The terms ξ_1 and ξ_2 represent random fluctuations in channel and source spaces respectively. The spatial covariances of ξ_1 and ξ_2 are mixtures of covariance components at each level. In sensor space, we assume $C_1 = \alpha^{-1} I_n$ to encode the covariance of sensor noise, where I_n is a n-by-n identity matrix. In sources space, we express it as a covariance basis form,

$$C_2 = \sum_{i=1}^{k} \gamma_i V_i, \tag{2}$$

where $\gamma \equiv [\gamma_1, \gamma_2, \ldots, \gamma_k]^T$ is a vector of k non-negative hyperparameters that control the relative contribution of each covariance basis matrix, V_i. Various candidate covariance basis can be adopted. For example, Minimum norm model (MNM) [6] has only one component $V = I$. LORETA [7], a very popular method in EEG source localization, has two covariance components $V = \{I, G\}$ [8], where $G = 2\exp(A)$ models anatomic coherent sources, and G is the Green function of an adjacency matrix, A, with $A_{ij}\square[0, 1]$ encoding the neighboring relationships among nodes of the cortical mesh defining the solution space [8]. Multiple sparse prior (MSP) employed components, $V = \{q_1 q_1^T, q_2 q_2^T, \ldots, q_k q_k^T\}$ to describe activities in k patterns, where q_j is evenly sampled column of matrix G [5].

All of the above priors are based on spatial adjacent relation of sources. The spatial patterns extracted from fMRI have never been considered in EEG source localization. Here, we employed ICA to decompose fMRI data to get the different spatial patterns. The intensity values in each pattern are scaled to z scores and voxels with absolute z scores >3 were considered to be activated. V_i for ith pattern is:

$V_i = \sum_{j \in U} q_j q_j^T$, where U is a set of activated voxels for ith pattern. The educed NESOI utilizes the patterns to improve localization accuracy. Based on the priors, NESOI is solved by restricted maximum likelihood (ReML) [5].

3 Result

In our simulation, the temporal and spatial patterns were generated by performing ICA decomposition on simultaneous EEG/fMRI recordings of a patient with epilepsy. These 20 spatial patterns are presented in Fig. 1. In each simulation, five spatial patterns θ_S were selected at random and each has an independent temporal pattern θ_T to generate the simulated signal in sensor space, $Y = L\theta_S\theta_T$. The signal and white noise were mixed to simulate actual data. We repeated the simulation 256 times.

We employed two metrics to compare NESOI with MNM, LORETA and MSP. Spatial accuracy measures the distance between the most active dipole and the dipole with the largest conditional expectation. Temporal accuracy is the squared correlation between the true time course of the most active dipole and the conditional estimate from the dipole used to assess spatial accuracy [5] (Fig. 2).

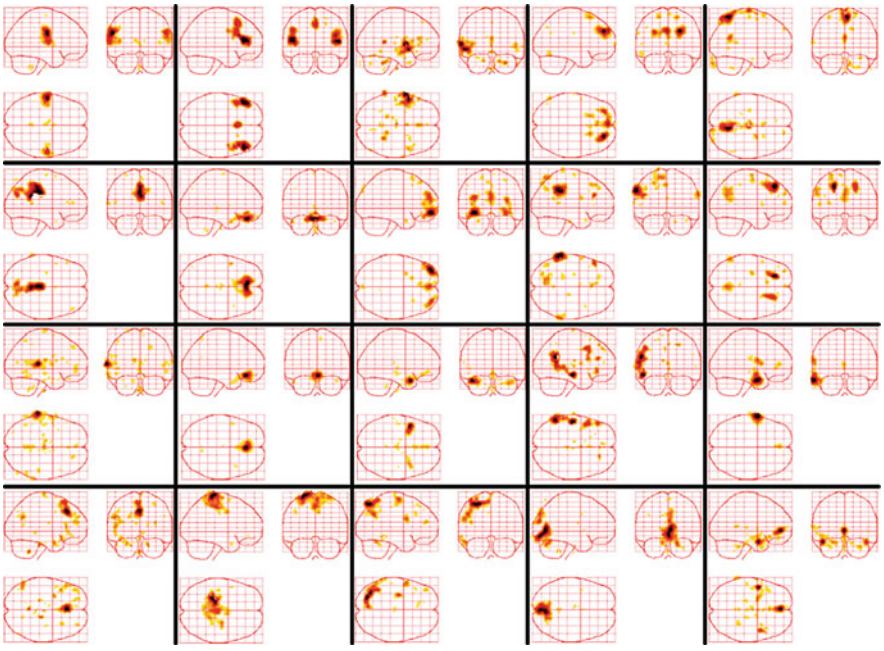


Fig. 1 The 20 spatial patterns. Five of these patterns were selected at random with temporal patterns to generate the simulated signal in sensor space (not shown)

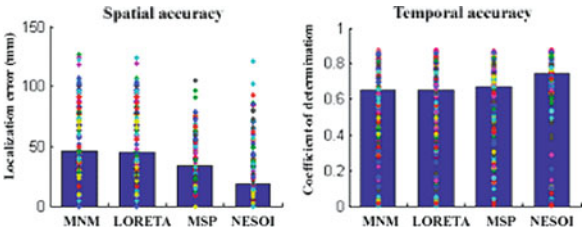


Fig. 2 Spatial accuracy (*left panel*) and temporal accuracy (*right panel*) measures for the four models under each of the 256 realizations. The *bars* represent the mean value and the *dot* is the result of each realization

It is clear that, both in terms of spatial and temporal accuracy, the NESOI model is substantially better than any one of the others. It is also interesting to note that MSP is better than the classical MNM or LORETA, and this difference is much bigger. This result is consistent with the previous report [5]. We should emphasize that only three covariance components in NESOI are derived from the patterns used in the simulation step and the others are inaccurate priors. However, with automatic selection the reverence prior, the solution was not influenced by the inaccurate prior when accurate and inaccurate location priors were used simultaneously.

4 Discussions

In this paper, we have described a new approach, NESOI, to employ fMRI to the distributed source reconstruction problem in EEG. Our key contribution is the employment of multiply fMRI spatial patterns in the EB model. With ReML algorithm, the reverence spatial patterns are automatically selected and localization accuracy is distinctly improved. Our computer simulation has demonstrated the validity and effectiveness of the method.

Comparing with traditional fMRI-constrained EEG inversion, NESOI allows the user to be agnostic with the exact form of the fMRI response and to link the fMRI coherent fluctuation networks with EEG feature. As brain networks established from fMRI are always disturbed by low temporal resolution, NESOI can be applied to provide more temporal details of these networks.

Acknowledgments This project was funded by grants from the National Nature Science Foundation of China (# 30525030, 60736029, 30870655).

References

1. Dale, A.M., Halgren, E.: Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. Curr. Opin. Neurobiol. **11** (2001) 202–208.
2. Liu, A.K., Belliveau, J.W., Dale A.M.: Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. Proc. Natl. Acad. Sci. **95** (1998) 8945–8950.

3. Phillips, C., Rugg, M.D., Friston, K.J.: Anatomically informed basis functions for EEG source localization: combining functional and anatomical constraints. Neuroimage. **16** (2002) 678–695.
4. Jacobs, J., Hawco, C., Kobayashi, E., Boor, R., LeVan, P., Stephani, U., Siniatchkin, M., Gotman, J.: Variability of the hemodynamic response as a function of age and frequency of epileptic discharge in children with epilepsy. Neuroimage. **40** (2008) 601–614.
5. Friston, K.J., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J.: Multiple sparse priors for the MEG/EEG inverse problem. Neuroimage. **39** (2008) 1104–1120.
6. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems. New York: John Wiley (1977).
7. Pascual-Marqui, R.D.: Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. **24** (2002) 5–12.
8. Harrison, L.M., Penny, W., Ashburner, J., Trujillo-Barreto, N., Friston, K.J.: Diffusion-based spatial priors for imaging. Neuroimage. **38** (2007) 677–695.

ICA Analysis for Decision-Making Task on fMRI

Xujun Duan, Huafu Chen, Zhengyong Pan, Wei Liao, Danhui Wang, and Jurong Ding

Abstract Brain functional activation of decision-making on fMRI experiment is seldom analyzed by data driven method to reveal the distinct brain functional network of decision-making. In this paper, a spatial independent component analysis (sICA) is presented to detect the functional network underlying decision-making by analyzing the unmixing matrices temporal components corresponding to spatial separated pattern. Our results show that ICA could separate various sub-functional activation networks to understand the neural mechanisms underlying decision-making, which mainly include functional network of conflict detection, cognitive and emotion.

Keywords fMRI · Decision-making · Iterated Prisoner's Dilemma game · sICA

1 Introduction

In recent years, there has been increasing interest in understanding the neural mechanisms underlying decision-making [1–3]. It is key issue that to gain brain functional network for different decision-making pattern by fMRI data process.

Among all the multivariate approaches, independent component analysis (ICA) is the predominance recently and is a valuable tool for the multivariate data-driven analysis of fMRI data [4, 5] for cognitive experiment or resting state networks [6]. However, ICA is not applied to analyze decision-making task. In the present study, sICA method is presented to separate the different activation patterns of the complex decision-making process by analyzing the temporal information of unmixing matrix corresponding to spatial component patterns.

H. Chen (✉)
Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China
e-mail: chenhf@uestc.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_62,

2 Material and Methods

2.1 Spatial ICA Model

In sICA, fMRI data sets were assumed to be a linear mixture of spatially independent component patterns, including background noise and patterns of brain activation signals [4], thus, the signal model of the sICA is presented as:

$$X = AS \tag{1}$$

where, X is a T-by-M (T is the number of time points and M is the number of voxels of the volume) observed fMRI signal, S is an N-by-M matrix whose rows S_i ($i = 1, ..., N$) contain the spatial process ($N \leq T$) and A is a T-by-N unknown time-invariant mixing whose columns A_j ($j = 1, ..., N$) contain the time course of the N processes. The separated spatial pattern matrix estimated is then written as:

$$S = A^{-1}X = WX \tag{2}$$

where, W is an N-by-T unmixing matrix. The aim of the sICA is to find the linear unmixing matrix, W and the separated source matrix, S, and then to get the original mutually independent signals or patterns S. In this study, the fast-fixed ICA algorithm was used for the sICA to estimate the matrix, W and matrix, S [7]. In the present work, the pre-processed of fMRI data were first reduced to 40 dimensions using PCA, in which >90% of the variability in the data was retained [8]. Using the sICA, data can be decomposed into 40 spatially independent components and corresponding temporally independent components (IC). For each IC, the waveform corresponds to the time course (TC) of a specific pattern of brain activity, and the intensity with which this activity is found across the voxels is expressed in the associated spatial map.

2.2 Decision-Making Task

We investigated event-related neural activations with fMRI as one subject played the iterated Prisoner's Dilemma Game [9] in the scanner with his partner outside the scanner. In this game, two players independently choose to either cooperate with each other or not, and each is awarded a sum of money that depends upon the interaction of both players' choices in that round. There are four possible outcomes of each around: player A (scanned) and player B (outside) cooperate (CC), player A cooperates and B defects (CD), A defects and B cooperates (DC), A and B defect (DD). These outcomes would be revealed to both players in the end of each round. Each cell of the payoff matrix (Fig. 1) corresponds to a different outcome of a social interaction [9]. This Prisoner's Dilemma game was consisted of 20 rounds, with each round lasting 19 s.

Fig. 1 The payoff matrix. Dollar amounts in parentheses are awarded to player B

		Player A (Scanned)	
		Coop	Defect
Player (B)	Coop	¥5 (5)	¥6 (0)
	Defect	¥0 (6)	¥1 (1)

2.3 Data Acquisition and Analysis

Functional imaging data was acquired on a 3.0 T scanner, GE-Signa (Huaxi MR Research Center, Chengdu, China). The gradient echo EPI sequence parameters were as follows: 30 transverse slices, TR = 2000 ms, TE = 30 ms, FOV = 24 cm, matrix = 64 × 64, voxel size = 3.75 × 3.75 × 4 mm^3 (without gap), flip angle = 90°. We collected 164 volumes per subject. The first 4 volumes for scanner calibration and for the subjects to get used to the circumstances were discarded before data processing, and left 160 volumes for analysis. The functional image data were preprocessed, and analyzed using the SPM2 software package (www.fil.iom.ucl.uk/spm).

3 Results

sICA was conducted for fMRI data using the fast ICA algorithm in GIFT software (http://icatb.sourceforge.net/, version 1.3b). Data were decomposed into 40 components. The voxels with Z values ($|Z| > 5.0$, $p < 0.001$) [4, 5] were in the light of the active voxels and overlaid the anatomical image. Figure 2 showed a separated spatial component patterns and its corresponding time course under the experimental

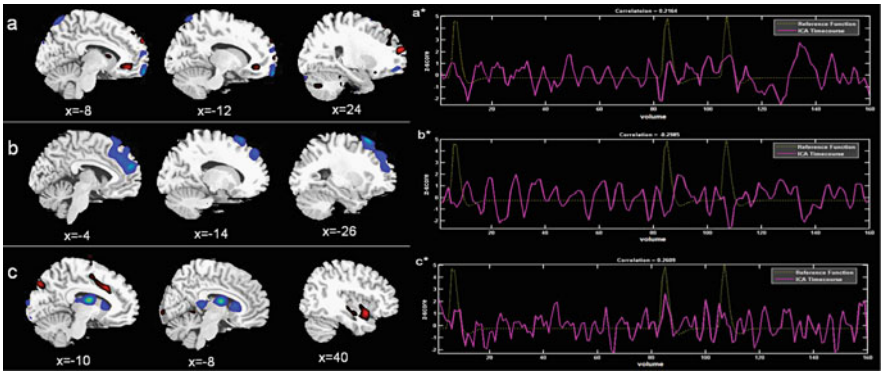


Fig. 2 **a**–**c**) Brain activated functional network of three spatial components for CC outcomes. (**a***–**c***) IC time course and reference function of experiment paradigm

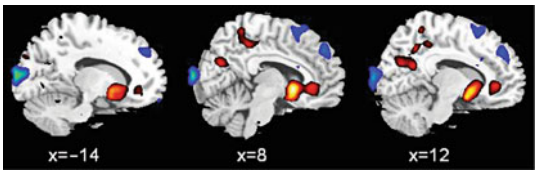


Fig. 3 SPM results for CC outcomes

paradigm, Fig. 2a–c are the three spatial separated components, while Fig. 2a*–c* are IC time course (red solid line) which highly correlated with the experiment paradigm's reference time course (yellow dotted line). As shown in Fig. 2, the first component represents functional network of conflict detection which include ACC; the second component represents functional network of cognitive process which include middle frontal gyrus, superior frontal gyrus and medial frontal gyrus; and the third component represents functional network of emotion which include insula, caudate nuleus and thalamus [10].

To test the effect of sICA method for detecting the brain functional network of decision-making, we compared our sICA results to SPM results. SPM imaging results for CC condition was shown in Fig. 3. The activation areas included precuneus, middle OFC, ACC, caudate, putamen and the deactivation areas included superior frontal gyrus and medial frontal gyrus. These areas were composite of a functional network which was almost the same as the combination of the three separated component in Fig. 2 decomposed by ICA, indicating that ICA method can detect brain functional activation of decision-making as same as SPM.

4 Conclusion

In sum, in this study, we focused on the ICA application in the decision-making fMRI data processing. Our result shows that ICA can provide more brain functional network information to understand the neural mechanisms underlying decision-making, and it's a reasonable method to reveal the complex cognitive process of decision-making.

Acknowledgments This work was supported by NSFC, 9082006 and 30770590, Key research project of science and technology of MOE (107097) and 863 Program No: 2008AA02Z4080.

References

1. Sanfey, G.: Social decision-making: insights form game theory and neuroscience. Science. **318** (2007) 598–602.
2. Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., Hanakawa, T.: Functional activity related to risk anticipation during performance of the low a gambling task. Neuroimage. **24** (2005) 253–259.
3. Krain, L., Wilson, A.M., Arbuckle, R., Castellanos, F.X., Milham, M.P.: Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. Neuroimage. **32** (2006) 477–484.

4. McKeown, M.J., Jung, T.P., Makeig, S., Brown, G., Kindermann, S.S., Lee, T.W., Sejnowski, T.J.: Spatial independent activity patterns in functional MRI data during the stroop color-naming task. Proc. Natl. Acad. Sci. USA. **95** (1998) 803–810.
5. McKeown, M.J., Hansen, L.K., Sejnowski, T.J.: Independent component analysis of functional MRI: what is signal and what is noise?. Curr. Opin. Neurbiol. **13** (2003) 620–629.
6. Mantini, D., Perrucci, M.G., Del Gratta, C., Romani, G.L., Corbetta, M.: Electrophysiological signatures of resting state networks in the human brain. Proc. Natl. Acad. Sci. **104** (2007) 13170–13175.
7. Chen, H., Yao, D.: Discussion on the choice of separated component in fMRI data analysis by spatial independent analysis. Magn. Reson. Imaging. **22** (2004) 827–833.
8. Chen, H., Yao, D., Chen, W., Chen, L.: Delay correlation subspace decomposition algorithm and its application in fMRI. IEEE Trans. Med. Imaging. **24** (2005) 1647–1650.
9. Rilling, J.K., Gutman, D.A.: A neural basis for social copperation. Neuron. **35** (2002) 395–405.
10. Sanfey, G., Riling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D.: The neural basis of economic decision-making in the ultimatum game. Science. **300** (2003) 1755–1758.

Single Trial Extraction of Cognitive Evoked Potential by Using EMD Based Method

Cailin Tao, Ling Zou, Yunfeng Li, and Zhenghua Ma

Abstract Empirical mode decomposition (EMD) is a new signal decomposition method, which could decompose the non-stationary signal into several single-component intrinsic mode functions (IMFs) and each IMF has some physical meaning. This paper studied the single trial extraction of cognitive evoked potential by combining EMD and wavelet threshold filter. Experimental results showed that the EMD based method can separate the noise out of the event related potentials (ERPs) and effectively extract the weak ERPs in strong background noise. It manifested as the waveform characteristics and root mean square error (RMSE).

Keywords EMD · Wavelet threshold · ERP · Single trial extraction

1 Introduction

Evoked related potentials (ERPs) have become very useful for cognitive study. The most common way to visualize the ERPs is to take an average over time locked single-trial measurements. But this method ignores the variation from trial to trial in the task-related cognitive process. Thus, the goal in the analysis of ERPs is the estimation of the single potentials that we call single-trial extraction. Several techniques have been proposed to improve the visualization of the ERPs from the background EEG with various successes [1–2].

Empirical mode decomposition (EMD) is a promising one for non-linear and non-stationary time series [3]. The fluctuations in different scales are stepwise disaggregated and a series of data sequence with different characteristic scales are produced. Each sequence is known as the intrinsic mode functions (IMFs). The

L. Zou (✉)
Faculty of Information Science & Engineering, Changhzou University, 213164; State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 100875 Beijing, China
e-mail: zoulingme@yahoo.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_63,

local zero-means of IMFs are symmetric and IMFs contain the same numbers of zero-crossing point and the extreme values. The lowest frequency component of the IMFs is on behalf of the signal trend. In this paper, ERPs were decomposed by the EMD and the IMFs were chose by energy distributions after filtered by the wavelet threshold filter, the single trial ERPs was reconstructed finally.

2 EMD-Based Method

EMD decomposition method can effectively extract the trend of a data sequence from the data sequence. It works like an adaptive high pass filter. It sifts out the fastest changing component of a composite signal first [4]. The sifting process is:

(1) For ERPs signal x(t); m(t) is the envelopes mean.
(2) Extract the detail h(t)=x(t)–m(t) and regard h(t) as new $x_1(t)=h(t)$; $h_1(t)=x_1(t)-m_1(t)$;. . .;$h_k(t)=x_k(t)-m_k(t)$, until two consecutive sifting results between 0.2 and 0.3 [5], then $imf_1(t)=h(t)$, $r_1(t)=x(t)-h(t)$.
(3) Let the residual $r_1(t)$ be a new signal, $x(t)=r_1(t)$.
(4) Repeat (2), (3), until there are no more than two extremes of the residual, obtain $imf_2, \ldots, imf_j, \ldots, imf_n$.
(5) x(t) is finally written as the sum of mode time series $imf_j(t)$ and the residual $r_n(t)$:

$$x(t) = \sum_{j=1}^{n} imf_j(t) + r_n(t). \tag{1}$$

(6) The IMF components were filtered by wavelet threshold method with an improved semi-soft threshold function, which could overcome the shortcomings of hard and soft threshold function [6].

$$\eta(w,T) = \begin{cases} w + T - T/(2k+1) & w < -T \\ w^{2k+1}/\left[(2k+1)T^{2k}\right] & |w| \le T \\ w - T + t/(2k+1) & w > T \end{cases}. \tag{2}$$

Where w=imfj is taken as the wavelet coefficient. $m = 1, \ldots, 2^{j-1}$, $\sigma = median\left(\left|imf_{j-1,m}\right|\right)/0.6745$; $T = \sigma\sqrt{2\log N}/\log(j+1)'$ is the threshold. When k tends to infinity, it is the standard soft threshold function. And there defines k = 5.

Suitable wavelet basis and decomposition scale should be selected according to the IMF signals. By testing, we chose "db4" with 7 levels. IMFs' selection was according to the practical energy distributions (refer to the following discuss). Finally the signal was reconstructed [7].

3 Discuss

EEG data were obtained from online free EEG/ERP database [8]. In this experiment, the subject covertly attended to the selected location on the computer screen responded with a quick thumb button press only when a square was presented at this location. This dataset contained only targets stimuli presented at the two left-visual-field attended locations for a single subject of 80 trials. EEG activity were recorded by 32 electrodes on the scalp and digitized at 128 Hz. For each trial, 3 s of data was saved from the pre-stimulated –1,000 to 2,000 ms.

We selected 10 trials of FZ, F3 and F4 electrodes randomly and analyzed them by EMD-based method. The major energy of ERPs in this experiment was at 300 ms which was the basic principle for the IMFs selection. Statistics showed that the signal energy of P300 in the experimental was mainly focused on the imf_4. Figure 1a showed the results at F3 randomly selected from some trail (here, we chose trial 6). Figure (1)–(6) was the energy distributions of imf_1 to imf_6. After rejecting imf_1, imf_2, imf_5, imf_6, the reconstructed signal (8) rejected huge fluctuation, kept characteristics and was more smoothness compared to without rejection (7).

Statistical analysis of the 10 trials at the three electrodes showed that the latency ($F(2, 27) = 0.232$, $P = 0.794 > 0.05$) and amplitude ($F(2, 27) = 0.610$, $P = 0.550 > 0.05$) of N100, the latency ($F(2, 27) = 1.665$, $P = 0.208 > 0.05$) and amplitude ($F(2, 27) = 1.795$, $P = 0.186 > 0.05$) of P300 were no significant difference. Single trial of extracted ERPs at the point of FZ, F3 and F4 were shown in Fig. 1b. The waveforms by using EMD method (bold lines) and superposed average were nearly superimposed.

The differences between EMD based and averaged method were not significant, see Table 1. The superposed average of 80 trials was taken as standard ERPs for calculating the RMSE of extracted ERP. And the averages of RMSE of extracted ERPs are shown in Table 1.

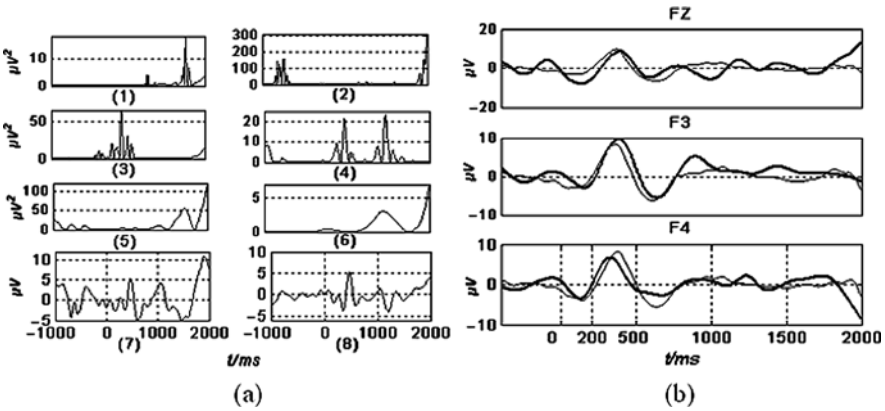


Fig. 1 (**a**) The energy distributions and choice of IMFs; (**b**) Results comparison of EMD method and superposed average

Table 1 Differences of features between the single trial data by EMD and superposed averaged

Point	Feature	Method	Latency (ms)	Amplitude (μV)	Average of RMSE
FZ	N1	EMD	125	16.8686	4.219
		Averaged	109	12.7630	
	P3	EMD	383	13.6489	
		Averaged	359	15.9479	
F3	N1	EMD	148	12.7821	4.212
		Averaged	50	11.4212	
	P3	EMD	383	15.3447	
		Averaged	359	14.7568	
F4	N1	EMD	125	10.1450	2.474
		Averaged	133	12.2223	
	P3	EMD	320	9.0932	
		Averaged	367	13.8710	

4 Concluding Remarks

After decomposition by EMD method, the instantaneous frequency is from high to low. But for the same component, the instantaneous frequency at different time can be varied considerably. In this paper, the de-noising effects of EMD based method are the same with the superposed average. It is mainly manifested as the RMSE and the characteristics of waveforms. Statistical analysis showed that the there was no significant difference among the 10 trials data. EMD could be further researched in the application of biological signal analysis.

Acknowledgments This work was supported by the open project of the State Key Laboratory of Cognitive Neuroscience and Learning, the open project of the Beijing Key Lab of Applied Experimental Psychology at the Beijing Normal University and QingLan Project of Jiangsu Province.

References

1. Dien, J., Beal, D.J., Berg, P. Optimizing principal components analysis of event-related potentials matrix type, factor loading weighting, extraction, and rotations. Clin. Neurophysiol. **116** (2005) 1808–1825.
2. Zou, L., Zhou, R.L., Hu, S.Q., et al. Single trial evoked potentials study during an emotional processing based on wavelet transform. Proceedings of the 5th International Symposium on Neural Networks, Vol. 5263 (2008) 1–10.
3. Huang, N. Shen, Z. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non stationary time series analysis. Proc R Soc London **454** (1998) 903–995.
4. Flandrin, P., Rilling, G., Goncalves, P. Empirical mode decomposition as a filter bank. IEEE Signal Process Lett **11** (2004) 112–114.
5. Huang, N.E., Shen, Z., Long, S.R., et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time-series analysis. Proc R Soc Lond A **454** (1998) 903–995.

6. Zhang, X.P., Desai, M.D. Adaptive de-noising based on SURE risk. IEEE signal Process Lett **5** (1998) 265–367.
7. Wang, M., Li, B.C. An improved speech enhancement method based on empirical mode decomposition and wavelet transform. Signal Process **24** (2008) 237–241.
8. http://sccn.ucsd.edu/eeglab/downloadtoolbox.html

Multiple Neural Networks Involved in Chewing Gum: A Functional MRI Study Using Group-ICA

Fan Li, Hui Shen, Baojuan Li, Xiaoling Zhou, and Dewen Hu

Abstract Mastication is an essential physiological function of the human neural system. Previous studies using univariate approach have found significant increases associated with chewing in the BOLD signal in the widespread brain regions. However, mechanism of functional organization across these regions and any relationship to cognitive processing are not clear at present. In this paper, we hypothesized that there exist multiple functional networks underpinning a chewing gum task, these networks being engaged in different brain function, such as sensorimotor and cognition. We used group independent component analysis (Group-ICA) to separate these networks based on functional magnetic resonance imaging (fMRI) data acquired during a chew-gum block task. Our results demonstrated that three significant neural networks identified by independent components (ICs), corresponding to sensorimotor, cognitive and emotional, and syntax function, involved in the chewing gum task. This study also provides further evidence for the hypothesis that there exist neural circuits associated with higher cognitive information processing in chewing-gum movement.

Keywords Functional network · Group independent component analysis (Group-ICA) · Chewing gum · Neuroimaging

1 Introduction

Mastication is an essential physiological function of the human neural system. In the early literature, PET studies reported increased blood flow in the primary sensorimotor cortex (SMC), supplementary motor areas (SMA), cerebellum and striatum during chewing task [1]. Recent fMRI study found chewing-related brain regions

D. Hu (✉)
College of Mechatronics and Automation, National University of Defense Technology, Changsha 410073, China
e-mail: dwhu@nudt.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_64,

include the SMC (extending down into the upper bank of the operculum and insula), SMA (extending down into the cingulated gyrus), thalamus, and cerebellum [2]. In the present work, we concern the mechanism of functional organization and information processing in the large-scale level of whole brain in mastication. Our goal is to apply Group-ICA into identifying the salient large-scale networks of the human brain during chewing gum and to search possible positive effects of chewing. We hypothesized that there are several functional networks underpinning the chewing gum process, one of them involved in sensorimotor function and others responsible for some different cognitive activities. We also expected to give some physiological explanation for cognitive effects of chewing gum based on these separated multiple functional networks during chewing gum tasks.

2 Materials and Method

2.1 Subjects and Task Paradigm

Sixty neurologically healthy college volunteers between the ages of 18 and 35 years were recruited by advertisement on BBS of the college. Twenty eight females and 32 males participated in this study. All the subjects had normal mastication function, and none were taking medication, abusing alcohol or illicit drugs. The subjects were instructed to minimize head movements during jaw movement, and data from participants where the heads were evaluated to have moved more than 0.75 mm would be discarded. Finally, the data of 38 participants were included in the analysis. Research consent form was obtained from each subject before the experiment. This study was approved by the Institutional Review Board of Beijing MRI Center for Brain Research.

Each participant performed the following two tasks: chewing gum and rest. The experiment were designed in a block manner (eight task blocks and eight rest blocks, each block of 25 s duration, alternated for a total scanning times of 400 s). The scanner was in the acquisition mode for 8 s before each series to achieve steady-state transverse magnetization. All performed in both parts of the experiment, having been instructed and given practice out of the scanner. In the gum chewing task, all participants selected their favorite chewing gum prior to the test. The participants were instructed to chew naturally and constantly throughout the chewing gum session.

2.2 Data Acquisition

The fMRI experiments were performed on a 3.0T Siemens Trio system in Beijing MRI Center for Brain Research. The head was positioned in a standard head-coil and fixed with cushions. All participants were instructed to minimize head movements and follow the instruction word (chew/rest) on the project screen. Earplugs were

provided to avoid auditory discomfort noises during the measurements. Prior to the fMRI examination, high-resolution, sagittal T1-weighted magnetic resonance (MR) images were obtained to determine the imaging plane parallel to the AC (anterior commissures)-PC (posterior commissures) line. The functional MR image data were acquired by the gradient-echo echo-planar imaging (EPI) sequence. An EPI pulse sequence was used with parameters as follows: Repetition time (TR)=3,000 ms, Echo time (TE)=30 ms, Flip angle(FA)=90°, Field of view (FOV)=240 mm, matrix=64 × 64, slice thickness=3.8 mm, gap=0.2 mm. Thirty two transversal slices of functional images covered the whole cortices and the cerebellum. It has been considered difficult to obtain sufficient fMRI data during jaw movement because the associated head motion creates artifacts on images. To avoid these artifacts, larger voxels ($3.75 \times 3.75 \times 4\ mm^3$) were used, allowing some head motion of the participants. Functional images with 135 volumes were acquired for this experiment.

2.3 Data Analysis

The procedure of data analysis mainly included two steps. The first step was data pre-processing including realignment, normalization, and smoothing, implemented by using SPM2 software. The second step was implemented by the Group-ICA program.

As a data-driven method, ICA has been used to separate fMRI data in spatial independent components together with their associated time courses [3]. Group-ICA is extended for multi-subject analysis from individual ICA and can be used to generate random-effect statistical inferences across subjects [4]. There are three main stages to Group-ICA: data compression, ICA separation and back-reconstruction. The outputs from these stages are multiple time courses. Each time course has an image map associated with it. The purpose of data compression is to reduce the size of the subject's data. In Group-ICA, Principal Components Analysis is used to reduce the dimensions. The fist PCA is used to reduce the individual data. Then the subjects are concatenated into groups and perform the second PCA. The next step is to use ICA to find independent components in the reduced data-set. Group-ICA includes some ICA algorithms such as Infomax, FastICA, Jade. Here we chose Infomax since it maximizes the information transfer from the input to the output of a network using a non-linear function. The components resulting from ICA represent group components, so the back-reconstruction uses the aggregate components of ICA and the results from data reduction step to compute the individual subject components [5].

In the Group-ICA processing, the preprocessed data were input into GIFT. The optimal number of components was estimated by the Akaike's information criterion coefficient, and set to be 20 in our study. After performing PCA, the data were separated by ICA using the Infomax algorithm. The components resulting from ICA represented group components, and through the back-reconstruction step it could

find the individual subject components. We used the correlation coefficient between the mean time course of each IC and the task paradigm to determine the task-related ICs. Only the ICs with positive correlation coefficients above a designated threshold ($p < 0.0001$) were included in next analyses. Then one sample t-test (threshold: uncorrected, $p < 0.001$, $k = 20$ voxels) was used to ascertain each group result of task-related IC.

3 Results

We found that there were three task-related ICs, IC1, IC2 and IC3, separated from Group-ICA analysis. The correlation coefficients between the task paradigm and the time course of each of the three networks were, respectively, 0.876, 0.74, and 0.668.

Our results showed that the IC1 involved the following areas: the anterior lobe of bilateral cerebellum, midbrain, thalamus, the precentral gyrus (BA4/6), lentiform nucleus (putamen), insula, the medial frontal gyrus (BA6), claustrum and post-central gyrus (BA2). The IC2 included these regions: the bilateral cerebellum, the subcallosal gyrus (BA34), the anterior cingulated (BA25), caudate, the middle frontal gyrus (BA10), the superior frontal gyrus (BA9), the inferior parietal lobule (BA40), angular gyrus (BA39) and cingulate gyrus (BA24). The IC3's corresponding activated brain regions consisted of inferior frontal gyrus (BA47), claustrum, caudate and thalamus.

4 Discussions

As we predict, the functional network component IC1 mainly represented sensorimotor network of chewing-gum. The IC1 mainly involved BA4 and BA6 which played an important role in integrating sensory information and planning and executing movements [6]. The IC1 also involved the sensorimotor cerebellum, which distributed across the anterior lobe of the cerebellar hemispheres [7]. Other areas, such as putamen, thalamus, played an important role in assorting with movements [8]. In a conclusion, the distributed brain regions involved in the IC1 network were associated with sensorimotor function of chewing gum, and consistent with the results from many previous studies using univariate approaches.

It's interesting and significant that the second functional network belongs to a cognitive and emotional network. The IC2 involved the lateral cerebellar hemispheres, also called "cognitive cerebellum" [7]. For the anterior cingulate cortex (ACC), it seems to be especially involved when effort is needed to carry out a task such as in early learning and problem solving [9]. Many studies attribute functions such as error detection , anticipation of tasks, motivation, and modulation of emotional responses to the ACC [10]. The ACC area is also associated with many functions that require conscious experience, and better emotional awareness has a relationship with improved recognition of emotional cues or targets which

is reflected by ACC activation [11]. The caudate has been demonstrated that it is highly involved in learning and memory, particularly regarding feedback processing [12]. As for the angular gyrus, it is a region of the brain in the parietal lobe, and involved in a number of processes related to language and cognition [13]. Furthermore, the executive functions of the medial frontal gyrus involve the ability to recognize future consequences resulting from current actions, to choose between good and bad actions, therefore, it is involved in higher mental functions [14]. In fMRI experiments, Goldberg et al. have found evidence that the superior frontal gyrus is involved in self-awareness, in coordination with the action of the sensory system [15]. The inferior parietal lobule is suggested to be associated with decision making under uncertainty [16]. These results support recent finding of the fronto-parietal cognitive network involved in mastication [17]. On the other hand, as a part of the limbic system, the subcallosal gyrus and cingulate gyrus are involved with emotion formation and processing. This result might give some physiological explanation for current hypothesis that chewing-gum can make people relax and alleviate negative mood [18]. Chewing-gum significantly activated many brain regions associated with the cognition and emotion, indicating that people can benefit from this simple action. In this study, the significant cognitive component of functional network was detected by Group-ICA, which strongly supported the hypothesis that the neural circuit of higher cognitive information processing involved in gum-chewing movement.

By chance, we found the third functional network associated with a syntax function identifying by IC3. It included BA47, claustrum, caudate and thalamus. These areas have been suggested to have a relationship with syntax. This network may be caused by Chinese characters (subjects needed to follow the instruction word (chew/rest) on the project screen), and it was irrelevant to gum-chewing. The result suggested that ICA is a sensitive data-driven approach for the segregation of the relevant components.

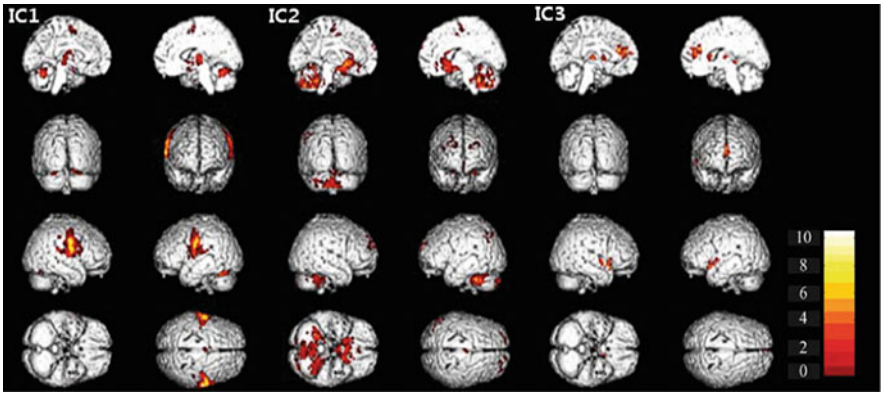


Fig. 1 The task-related ICs separated from Group-ICA analysis of gum-chewing (one sample t-test, height threshold: uncorrected, $p < 0.001$, $k = 20$ voxels)

5 Conclusions

To sum up, the results strongly support that multiple neural networks are involved in chewing gum. Besides the predicted sensorimotor functional network, a significant cognitive network is separated from fMRI data of chewing gum tasks, which provides further physiological evidence for recent hypothesis that chewing gum improves mood and reduces anxiety and stress.

Acknowledgments This work is supported by National Basic Research Program of China (2007CB311001), National Natural Science Foundation of China (60771062, 60835005, 90820304).

References

1. Baldonado, Momose, I., Nishikawa, J., Watanabe, T., Sasaki, Y., Senda, M., Kubota, K., Sato, Y., Funakoshi, M., Minakuchi, S.: Effector of mastication on regional cerebral blood flow in humans examined by positron-emission tomography with 150-labelled water and magnetic resonance imaging. Arch. Oral Biol. **42** (1997) 57–61.
2. Onozuka, M., Fujita, M., Watanabe, K., Hirano, Y., Niwa, M., Nishiyama, K., Saito, S.: Mapping brain region activity during chewing: A functional magnetic resonance imaging study. J. Dent. Res. **81** (2002) 743–746.
3. McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.-P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J.: Analysis of fMRI data by decomposition into independent spatial components. Hum. Brain Mapp. 6 (1998) 160–188.
4. Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J.: A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. 14 (2001) 140–151.
5. Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J.: Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp. 13 (2001) 43–53.
6. Blakemore, S.J., Frith, V.: The Learning Brain. Oxford: Blackwell Publishing. ISBN 1-4051-2401-6 (2005).
7. Dewen, H., Hui, S., Zongtan, Z.: Functional asymmetry in the cerebellum: A brief review. Cerebellum **7** (2008) 304–313.
8. Jones, E.G.: The Thalamus. Cambridge: Cambridge University Press (2007).
9. Allman, J.M., Hakeem, A., Erwin, J.M., Nimchinsky, E., Hof, P.: The anterior cingulate cortex. The evolution of an interface between emotion and cognition. Ann. N Y Acad. Sci. **935** (2001) 107–117.
10. Bush, G., Luu, P., Posner, MI.: Cognitive and emotional influences in anterior cingulate cortex . Trends Cogn. Sci. **4**(6) (2000) 215–222.
11. Lane, R.D., Reiman, E.M., Axelrod, B., Yun, L.S., Holmes, A., Schwartz, G.E.: Neural correlates of levels of emotional awareness. Evidence of an interaction between emotion and attention in the anterior cingulate cortex. J. Cogn. Neurosci. **10**(4) (1998) 525–535, 1998.
12. Steriade, M., Llinas, R.: The functional states of the thalamus and the associated neuronal interplay. Physiol. Rev. **68** (1998) 699–742.
13. Ramachandran, V.S.: A Brief Tour of Human Consciousness. New York, NY: Pi Press (2004) 60–82.
14. Kimberg, D.Y., Farah, M.J.: A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex, organized behavior. J. Exp. Psychol. Gen. **122**(4) (1993, 2004) 411–428.
15. Goldberg, I., Harel, M., Malach, R.: When the brain loses its self: prefrontal inactivation during sensorimotor processing. Neuron **50**(2) (2006) 329–339.

16. Vickery, T.J., Yuhong, V.J.: Inferior parietal lobule supports decision making under uncertainty in humans. Cereb. Cortex **19** (2009) 916–925.
17. Takada, T., Miyamoto, T.: A fronto-parietal network for chewing of gum: a study on human participants with functional magnetic resonance imaging. Neuroscience **360** (2004) 137–140.
18. Takahashi, T., Miyamoto, T., Terao, A., Yokoyama, A.: Cerebral activation related to the control of mastication during changes in food hardness. Neuroscience **145** (2007) 791–794.

Part VII
Sensory and Motor Dynamics

Dynamic Form Templates Determine Sensitivity to Biological Motion

Joachim Lange and Markus Lappe

Abstract Visual perception of biological motion shows a remarkable robustness against noise, fundamentally different from sensitivity to other moving stimuli. This is evidence for especialized mechanisms for biological motion perception that are more sensitive to biological motion than to other stimuli. Yet, the specifics of biological motion stimuli or the mechanisms which might explain the qualitative discrepancy between coherent motion and biological motion in terms of sensitivity remain elusive. In a combination of neurocomputational modeling and psychophysical experiments we investigated how form and motion signals influence sensitivity to biological motion in noise. With stimuli that vary in the amount of motion signals we tested the ability to detect and discriminate biological motion in human observers and in a dynamic neuro-cognitive model of biological motion perception. These results suggest that the sensitivity to human movements is caused by a specialization to the dynamic and complex pattern of the changing form of the body over time.

Keywords Vision · Visual motion · Visual form · Biological motion

1 Motion Integration for Simple and Biological Motion

Global motion perception is achieved by spatio-temporal integration of local motion signals. The sensitivity to simple translational motion is influenced by noise in such a way that the tolerance to the noise increases linearly with the number of stimulus dots [1, 2]. Point light displays of biological motion consist of stimulus dots that move smoothly during a walking sequence. For short time intervals the dot motion

M. Lappe (✉)
Department of Psychology, Institute II, Westf Wilhelms-University, Münster, Germany; Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, Westf Wilhelms-University, Münster, Germany
e-mail: mlappe@uni-muenster.de

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_65,

is approximately linear and provides a brief local motion signal. Theoretically, integration of these local motion signals over space and time might therefore result in a global impression of the stimulus, similar to the integration mechanisms proposed for simple translational motion. Unlike for simple coherent motion, however, the limited lifetime-technique, in which point motion is confined only to a limited number of succesive frames, showed that sensitivity to discriminate the walking direction of biological motion increased in a non-linear way, much stronger than expected from simple mechanisms that integrate motion over space and time [2].

2 Template Model for Biological Motion Perception

Our model uses a template-matching algorithm based on distance measures between stimulus dots and stick-figure templates [3, 4]. When tested with stimuli that were identical to the stimuli used experimentally, i.e. a limited lifetime-technique that lead to local motion signals in the stimuli [2], the increase in sensitivity in the model matched that of the published experimental data [3]. However, the model operates by simply analyzing the sparse form information in the point-light stimuli while ignoring the local motion signals. It predicts, therefore, that sensitivity to direction discrimination of biological motion should be independent of local motion signals. Confirmation of this prediction would corroborate the exceptionality of biological motion perception, and demonstrate that human perception of biological motion is not necessarily an adaptive mechanism but could be explained by simple template matching mechanisms.

A detailed description of the model can be found in [3]. Here, we give a short qualitative description of the two stages of the model. The first stage uses an algorithm that matches each stimulus frame independently to a set of stored static templates extracted from the walking sequences of nine persons. Half of the templates face to the left the others face to the right. By computing a distance measure between a specific stimulus frame and each of the template frames, the model extracts the template frame that matches best to the stimulus frame. This template is then used for a decision whether the stimulus frame shows a walker oriented to the left or to the right. Subsequently, this procedure is repeated for all stimulus frames and the model achieves an overall decision by averaging over all decisions obtained for the single frames. This stage, therefore, analyzes only the global form information available in the single stimulus frames. Subsequently, the information is forwarded to a second stage that analyzes the temporal order of the single stimulus frames, i.e. this stage analyzes the global motion information. We have shown that the output of the first stage correlates with behavioral results of human observers in discrimination task like the one presented in this study [3, 4]. Since we applied in the present study a direction discrimination task, only the first stage of the model is used.

3 Human Experiments and Model Simulations

The stimulus used for combined model simulation and experimental studies depicted a human body viewed from the side walking in place as on a treadmill [5]. It contained of a fixed number of dots that varied from 2 to 8. The positions of the dots were chosen randomly on the limbs [6]. The dots were either relocated to a new, again random position on the limbs after each frame ("lifetime 1") or they moved on this position with the limb for two frames before being relocated on a new location on the limbs ("lifetime 2"). The "lifetime 2" condition contained apparent motion signals, the "lifetime 1" condition did not. The stimulus was walking forwards and oriented either to the right or to the left, presented embedded in a field of random noise dots. The noise dots changed their position each frame to a new, randomly chosen position.

Four subjects (age 25–30) participated in the experiment. Their task was to indicate the orientation (left or right) of the stimulus. The stimulus was presented on a computer screen (refresh rate 100 Hz), covered a field of $5 \times 10°$, and consisted of white dots (5×5 pixels) on a black background. A single stimulus frame lasted 50 ms (five monitor frames). Stimulus velocity was 0.67 walking cycles/s. Stimuli lasted up to three walking cycles. The experiment was split into blocks of a constant number of stimulus dots and constant lifetime of the stimulus dots (1 or 2). In each block, a 2-up 1-down staircase determined the 70.7% correct noise threshold for discriminating the walking direction. Thresholds were plotted in a log-log-plot and fitted by linear regression.

The model simulations started by determining the recognition rate of 100 stimulus sequences with 2 dots per stimulus frame and a dot lifetime of 1. Subsequently, we added in a window (6.5×4 times the stimulus size) a fixed number of noise dots and again determined the recognition rate. We repeated the procedure until the recognition rates dropped to chance level. Then we plotted the recognition rates as a function of number of noise dots and fitted these data points with a sigmoid function in order to determine the noise thresholds for 70.7% correct responses. We repeated this procedure for all combinations of number of stimulus dots and lifetimes. The results were plotted in log-log-diagrams and fitted by linear regression.

Fig. 1 shows the noise thresholds measured for the four subjects (a–d) and the model (e) for both dot lifetimes. The model results are characterized by a linear increase in the log-log-plots for the noise thresholds for both lifetimes. The slopes of the linear regression are steeper than (1). Moreover, slopes for both conditions are virtually identical, i.e. the slopes did not show significant differences. In agreement with the model, all subjects revealed slopes that are steeper than 1 for dot lifetime (2). Furthermore, the data from the subjects confirmed the model predictions for lifetime 1: all slopes are steeper than 1 and all slopes are virtually identical to slopes obtained for dot lifetime 2, i.e. there are no significant differences in terms of slope steepness. All slopes are in the same range as the published data [2] for a walker with a dot lifetime of 2.

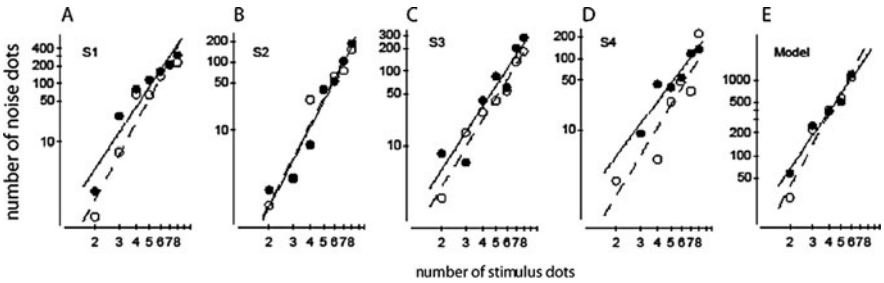


Fig. 1 Noise thresholds for 70.7% correct identification of the walking direction of the stimulus as a function of number of stimulus dots per stimulus frame. *Filled dots* represent data for conditions with a dot lifetime of 1, *open dots* for a lifetime of 2. Data are plotted in a Log–Log-diagram and fitted by linear regression (*solid line* for dot lifetime of 1, *dashed lines* for dot lifetime of 2). a–d show the psychophysical results for four subjects, e shows the data for the model simulations

4 Discussion and Conclusion

We tested our model prediction that local motion signals are irrelevant for sensitivity to direction discrimination of biological motion. First, we reproduced prior experimental results [2] in simulations of the form-based model. The linear regression for stimuli with a dot lifetime 2 revealed a slope of 3.18, thus confirming the nonlinear increase of sensitivity to biological motion stimuli. Second, the model predicted the same results for dot lifetimes of 1 as for dot lifetime 2 (slope of 2.53), i.e. that sensitivity to direction discrimination of biological motion should be unaffected by local motion signals. We confirm this prediction by showing that our subjects noise thresholds were virtually identical in both lifetime conditions. We conclude that our form-based model well describes the direction discrimination of biological motion.

In contrast to the direction discrimination tasks, for the task of *detection* in noise a linear relationship between number of stimulus dots and detection thresholds was shown for translational motion as well as for biological motion [2]. We suggest that detecting biological motion and discriminating its walking direction are two distinct processes which are not necessarily coupled. To detect biological motion in noise local motion signals may be important to segregate the stimulus from the noise [7, 3]. This process is linearly dependent on local motion signals but unspecific for biological motion. Our results suggest that a template-matching process is sufficient to achieve the second step of the recognition process, i.e. the discrimination of walking direction, which is specific for biological motion stimuli.

References

1. Barlow, H., Tripathy, S.: Correspondence noise and signal pooling in the detection of coherent visual motion. J. Neurosci. **17**(20) (1997) 7954–7966
2. Neri, P., Morrone, M.C., Burr, D.C.: Seeing biological motion. Nature **395** (1998) 894–896

3. Lange, J., Georg, K., Lappe, M.: Visual perception of biological motion by form: a template-matching analysis. J. Vis. **6**(8) (2006) 836–849
4. Lange, J., Lappe, M.: A model of biological motion perception from configural form cues. J. Neurosci. **26**(11) (March 2006) 2894–2906
5. Cutting, J.E.: A program to generate synthetic walkers as dynamic point-light displays. Behav. Res. Methods Instrum. Comput. **10**(1) (1978) 91–94
6. Beintema, J.A., Lappe, M.: Perception of biological motion without local image motion. Proc. Natl. Acad. Sci. USA **99**(8) (2002) 5661–5663
7. Cutting, J.E., Moore, C., Morrison, R.: Masking the motions of human gait. Percep. Psychophys. **44**(4) (1988) 339–347

Bistable Alternation of Point-Light Biological Motion

Marc H. E. de Lussanet and Markus Lappe

Abstract The facing-in-depth of point-light biological motion is ambiguous: the frontal and back view look the same. However, since earlier studies found a very strong perceptual bias in point-light biological motion, it is unknown whether it evokes an alternating (bistable) percept. In the present study, naive, untrained observers viewed point-light stimuli in half-profile view. All participants experienced spontaneous flipping of the orientation-in-depth, both for biological motion and necker cube displays. The number of perceptual flips was lower for the rocking cube than for the static one; and higher for biological motion than for rocking cubes. Contrary to earlier findings the participants did not have a perceptual bias. We conclude that ambiguous biological motion does evoke a bistable percept.

Keywords Biological motion · Depth perception · Perceptual bistability · Psychophysics

1 Introduction

Ambiguous visual stimuli are highly interesting for studying brain dynamics, because the brain is forced to arrive at a consistent interpretation, on the basis of the dynamical interaction of neuronal populations. Ambiguity may, e.g. lead to a simultaneous, conflicting percept (e.g. motion transparency), or to spontaneous spatial and temporal alternation of the conflicting percepts (e.g. necker cube or binocular rivalry). Biological motion displays human movements as moving point lights [1]. For a non-profile view, the facing-in-depth is ambiguous, i.e. there is no way to tell whether one sees the front or the backside. Biological motion is a particularly

M.H.E. de Lussanet (✉)
Allgemeine Psychologie, Psychologisches Institut II, Westfälische Wilhelms-Universität, Münster, Germany
e-mail: lussanet@uni-muenster.de

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_66,

interesting ambiguous stimulus, because it is a complex stimulus with very little information, and requires high-level visual processing [9].

Ambiguous displays of 3-D objects often evoke a bistable percept. The well-known necker-cube displays the ribbons of a cube without perspective deformation nor occlusion. If one looks at a necker cube for some time, the percept will be bistable or even multistable. That means that the percept will spontaneously change to a different possible percept, such as two possible 3-D interpretations, or even a flat wireframe [2]. Alternating percepts are thought to arise from the dynamical competition between neuronal cell populations that each represent possible, but mutually exclusive, percepts [3].

One could expect that the percept of point-light biological motion is bistable, like the necker cube. However, biological motion differs from the necker cube because it moves, and it differs from dynamic displays such as the rotating drum [4] in that its form changes dynamically. Thus, rhythmic biological motion might be special in that the perceptual dynamics interfere with the cyclic change in form. Moreover, observers have strong perceptual bias to see the frontal view [5, 6]. We investigate whether biological motion displays are bistable.[1] Since perceptual changes are linked to the amount of conflict in the stimulus [4], a strong perceptual bias could abolish the bistable percept of biological motion.

2 Methods

Stimuli were computed, using in-house developed programs, and were presented on a CRT monitor, 40 × 30 cm, 640 × 480 pixels, 180 Hz vertical refresh. Participants wore CrystalEyes-3 stereo goggles so the effective refresh rate was 90 Hz per eye. All stimuli consisted of 14′ red dots on a dark background.

Stimuli represented a repeated walking cycle without net displacement [7], or a cube (Fig. 1). The walker consisted of 14 dots, and was oriented 45° about the vertical. The cube consisted of 8 dots located on the corners. The cube was oriented 45° about the vertical and tilted 25°. When rocking, a sinusoidal tilt was applied with an amplitude of 10°. All stimuli were 3° wide. The movement period was always 1.4 s.

The experiment consisted of three 5-min sessions, respectively presenting biological motion, rocking cubes, and stationary cubes, repeated on 2 days. During each experimental session, the stimulus was continuously displayed. Every 30 s, an invisible, smooth transition to a different condition occurred (each condition occurred twice in each session). The condition of interest displayed an orthographic projection. Four other conditions that were not analyzed here, applied different depth cues to the stimulus. Each condition occurred twice during a session.

The nine participants did not know our goal, and were inexperienced with psychophysical experiments. The task was to report the current percept by a key-press

[1] Notice that we use bistability in the usual sense [3, 4], which is different from [2].

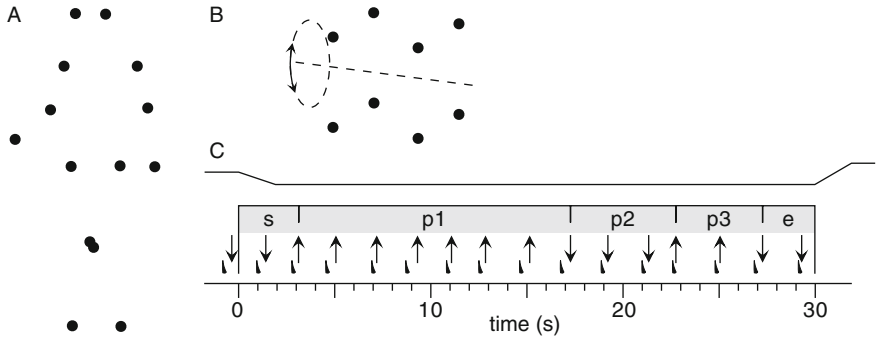


Fig. 1 The walking stimulus (**a**) and the Necker cube (**b**). **c** A 30-s interval of presentation. *Upper trace*: the presented condition with a transitional period on the first 2 s. *Triangles* depict the auditory cue to respond (*arrows*). In this example, there are three perceptual intervals: p1, p2, p3

on hearing a brief bell sound, on average every other second with a random jitter of 0.5 s (Fig. 1c; [8]). Valid responses were "Toward", "Flat", or "Away" for a walker or "From Above", "Flat", or "From Below" for a cube.

The perceptual intervals were selected from each 30-s presentation period. Only intervals were selected that started and ended with a perceptual change within this period. If no such interval was present, a 29-s interval was assumed. The cumulative distributions of response intervals were statistically compared using a bootstrap-version of the Kolmogorov-Smirnov test (ks.boot, R version 2.8), with 10^5 bootstraps.

3 Results

Figure 2 shows the cumulative perceptual intervals over all participants. The saturation of the distributions above 10 s shows that most intervals lasted less than 10 s. The jump at the last bin reflects that a number of 30-s periods consisted of just a single long perceptual interval that was interrupted at the end of the trial. This occurred most frequently in the rocking cube condition. All participants experienced at least one interval of 15 s or shorter in each condition. Thus, spontaneous perceptual changes occurred in all three conditions.

The cumulative distribution for biological motion lies in between those for the static and rocking Necker-cubes (Fig. 2). There were more short intervals for static cubes than for biomotion (ks.boot-$P < 0.0001$), and more short intervals for the latter than for rocking cubes (ks.boot-$P = 0.0001$). This shows that periodic motion reduces the number of short perceptual intervals, and thus the number of spontaneous perceptual changes. Importantly however, the result shows that the percept of point-light biological motion is bistable. The number of perceptual changes

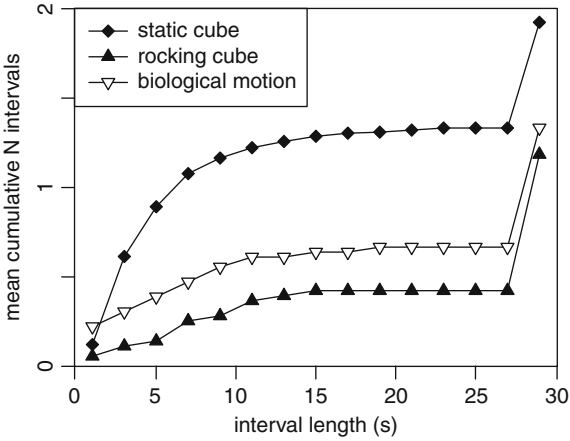


Fig. 2 Mean cumulative distributions of perceptual intervals. The last bin includes intervals that were interrupted by the end of the presentation period

was considerably lower than for static necker cubes, but more than in a rocking Necker-cube, which had a comparable amount of movement as the walking stimulus.

A "Flat" response was given just once for biological motion, but were frequent with static and moving cubes (10% of the intervals of <15 s). Remarkably, none of the participants showed a strong perceptual bias either for biological motion or for the static or rocking cube. Of the <15 s intervals 55% were "towards" and 43% were "away" responses for biological motion (43–49% for the necker cube). The same was found for the total viewing time: 53% of all responses were "towards" to 47% "away" for biological motion (49–45% for the necker cube).

4 Discussion

The present results show that point-light biological motion is a bistable stimulus, in the sense that the two possible percepts (frontal versus back view) alternate spontaneously when one looks at it for some time.

A remarkable finding of the present study is that the perceptual bias for a toward-percept was hardly present in our data, in contrast to published results [2, 6]. This apparent discrepancy probably is not due differences between the displays. We asked a number of naive observers, and orally they reported almost invariably a "toward" view of the walker. A possible explanation is that the perceptual bias is weak, and therefore easily suppressed; for example by depth cues such as linear perspective deformation, or occlusion. In the present study the participants did not have a forced choice but reported the depth percept over a longer period, which probably strongly reduced the bias.

Since ambiguous stimuli have multiple interpretations, the brain needs to arrive at a consistent interpretation. The cyclic motion that imposes the brain with a gradual

form change might interfere with this dynamic process. Indeed, the spontaneous perceptual changes in necker cubes were almost abolished by rocking motion in our population of inexperienced observers.

In biological motion the gradual, cyclic form change did not abolish spontaneous perceptual changes, but compared to the static necker cube, did reduce them. Since there was no perceptual bias, we find it unlikely that the biological motion stimulus suffered from reduced perceptual conflict [4]. Also, the form change did not abolish the bistability as in the necker cube. Instead, we propose that there is a neuronal conflict at a neuronal level that integrates the complete walking cycle, which is consistent with cell recordings in monkeys [9].

References

1. Johansson, G.: Visual perception of biological motion and a model for its analysis. Percept. Psychophys. **14** (1973) 201–211.
2. Kornmeier, J., Bach, M.: The necker cube–an ambiguous figure disambiguated in early visual processing. Vision Res. **45** (2005) 955–960.
3. Brouwer, G.J., Tong, F., Hagoort, P., van Ee, R.: Perceptual incongruence influences bistability and cortical activation. PLoS ONE **4** (2009) e5056.
4. Pastukhov, A., Braun, J.: A short-term memory of multi-stable perception. J. Vis. **8** (2008) 1–14.
5. Vanrie, J., Dekeyser, M., Verfaillie, K.: Bistability and biasing effects in the perception of ambiguous point-light walkers. Perception **33** (2004) 547–560.
6. Vanrie, J., Verfaillie, K.: Perceiving depth in point-light actions. Percept. Psychophys. **68** (2006) 601–612.
7. Lange, J., Lappe, M.: A model of biological motion perception from configural form cues. J. Neurosci. **26** (2006) 2894–2906.
8. Mamassian, P., Goutcher, R.: Temporal dynamics in bistable perception. J. Vis. **5** (2005) 361–375.
9. Oram, M.W., Perrett, D.I.: Responses of anterior superior temporal polysensory area (STPa) neurons to “biological motion” stimuli. J. Cogn. Neurosci. **6** (1994) 99–116.

Influence of Feedback Signals from Cat's Area 21a on Contrast Response Function of Area 17 Neurons

C. Wang, W. Burke, and B. Dreher

Abstract Feedback signals originating from cat's extrastriate visual cortices exert mostly excitatory influence on neuronal responses in area 17 (A17). In the present study we found that feedback signals from one of the "form-processing" extrastriate cortical areas, area 21a (A21a), modulate the contrast response function of A17 neurons. Thus, in a majority (13/18) of A17 cells examined there was a significant change in contrast response function during inactivation of A21a. While the magnitude of the responses to stimulus contrasts tends to decrease during inactivation of A21a there was no systematic change in the mid-contrast, i.e. the contrast at which the response of individual neurons reached 50% of its maximum response. Further analysis indicated that the effect on the magnitude of responses was stronger at high contrast. These results suggest that an overall excitatory feedback from A21a may play a role in enhancing contrast response in A17 neurons, especially at high contrast.

Keywords Contrast response function · Neurons · Feedback · Visual cortex · Cats

1 Introduction

Despite abundant feedback projections from "higher-order" to "lower-order" visual cortices little is known about their functional roles (see for reviews, [1, 2]). Previous studies based on single neuron recordings indicate that feedback connections exert primarily excitatory influence on responses of neurons in primary visual cortex [3–8]. Such an influence can also be detected by performing intrinsic signal optical imaging on primary visual cortices (areas 17 and/or 18, V1 and/or V2 [9, 10]) during transient inactivation of extrastriate visual cortical area, area 21a (A21a) or

C. Wang (✉)
Discipline of Anatomy and Histology, Bosch Institute, The University of Sydney, Sydney, NSW 2006, Australia
e-mial: chunw@anatomy.usyd.edu.au

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_67,

the postero-medial and postero-lateral suprasylvian areas of the cat. We postulated that the excitatory influence of feedback projections might also modulate some fundamental receptive field properties of neurons in primary visual cortex.

The dependence of the magnitude of response on stimulus contrast and contrast adaptation processes is a fundamental feature of vertebrate visual system and has been found already in "pre-cortical" stages of the mammalian visual pathways (retina: [11]; dorsal lateral geniculate nucleus (LGNd): [12, 13]). Although the contribution of the visual cortex to processing of contrast-related information is not yet fully understood the orientation-selective contrast adaptation observed in a substantial proportion of V1 and V2 cells in anaesthetised cats [14] and the attention-dependent increase in contrast sensitivity in area V4 neurons of awake macaque monkeys [15] suggests further modulation of contrast response function (CRF) in visual cortices. In the present study we examined in detail the influence of feedback projections from one of the "form-processing" higher-order cortical areas, A21a on contrast response function of area 17 (A17) neurons by reversibly inactivating ipsilateral A21a (by cooling it to 10°C). Our analysis suggests that feedback signals originating from A21a not only exert a significant, primarily excitatory, influence on the magnitude of the responses of A17 neurons but also in many cases induce substantial changes in CRF.

2 Results and Discussion

In the present study a total of 25 cells recorded extracellularly in the part of A17 visuotopically corresponding to the part of the visual field represented in A21a, abbreviated as A17(c21a) [5] were examined for CRF prior to, during and after reversible inactivation of A21a. Due to a significant variation between the control runs or non-monophasic CRF profiles, seven cells had to be excluded from further analysis. After determining the optimal stimulus orientation, direction and velocity we then examined CRF by presenting a series of optimally oriented light bars of various contrasts in interleaving fashion.

In a majority of A17 cells (13/18, 72%) inactivation of A21a resulted in significant changes to their CRFs ($0.001 < P < 0.016$, $n = 13$, F test). Figure 1 shows the model functions (Fig. 1a, d) used and examples of five A17 cells (Fig. 1b, c, e–g) whose CRFs were significantly affected by reversible inactivation of A21a. As illustrated in Fig. 1b, c the CRFs of both cells were well fitted to hyperbolic ratio functions. However, the effects observed in these two cells were very different. Thus, during inactivation of A21a the cell whose CRF functions are illustrated in Fig. 1b showed a profound reduction (51%) in the magnitude of its maximum response with almost no change (<4%) in the C_{50}, therefore mimicking a change in response gain control. On the other hand, during cooling of A21a the CRF of the other cell (Fig. 1c) shifted substantially to the right along the axis of stimulus contrasts and the C_{50} increased nearly 5-fold from 1.65 prior to cooling to 9.74. When compared to the scale of the change in the C_{50}, the increase in the magnitude of

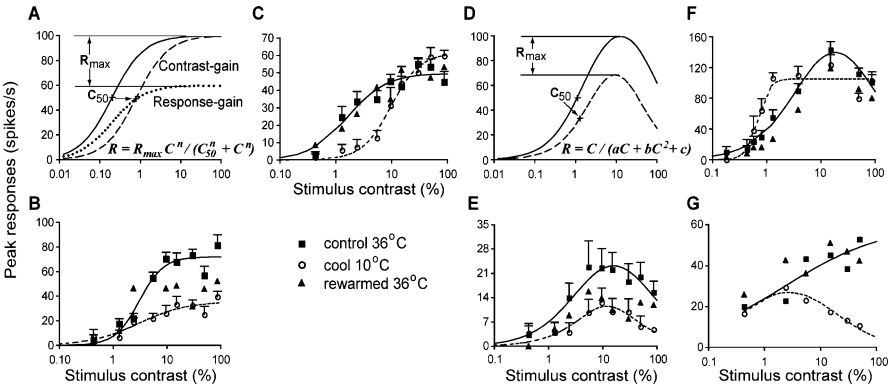


Fig. 1 Examples of model functions for contrast response functions (CRFs) and the effect of inactivation of A21a on CRFs of five A17 cells. (**a**) Variations of the sigmoid model functions are shown with *dotted* and *dashed lines* indicating an exclusive change in R_{max} (response-gain) and in C_{50} (contrast-gain) respectively in relation to an original function (*solid line*). (**b**, **c**) Two A17 cells showed a 51% reduction in response (**b**) and a predominant shift in the C_{50} (**c**) during inactivation of A21a. (**d**) A model function based on polynomial ratio. (**e**) An A17 cell whose CRF well fitted to a polynomial ratio function showed a substantial reduction (51%) in the maximal response and a small change (20%) in C_{50} during inactivation. (**f**) An A17 cell showing a clear response reduction at high contrasts exhibited clear facilitation at a number of low contrasts and the responses saturated much earlier during inactivation. (**g**) An example of a reduction in the dynamic range of contrast response function during inactivation of A21a. Contrast (**c**) = $(L_o - L_b)/L_b$, where L_o is the luminance of object bars and L_b is the background luminance, which was 1.0 cd/m^2 in the present study

the maximum response for this cell was relatively small (23%, from 49.8 to 61.6 spikes/s), thus indicating a prominent effect on a contrast gain control (cf. Fig. 1a).

In some cells we found, however, that the magnitude of the responses was considerably reduced at high contrasts, i.e. a bell-shaped CRF. A polynomial ratio function was used to fit those CRFs (Fig. 1d, see also examples in Fig. 1e, f). In most cases, the maximum response of the CRF was substantially reduced with a minor change in the C_{50} and the bell-shaped CRF was retained during inactivation of A21a as shown in Fig. 1e. On the other hand, the shape of the CRF may also be altered in addition to the magnitude of response and C_{50} during cooling as illustrated in Fig. 1f, g. Thus, these results suggest that the feedback signals from A21a can exert a strong influence on the CRFs of A17 cells (Fig. 1b, c, e–g) and, as a result, the maximum response and/or the C_{50} are affected.

An average *reduction* of 31% (range 11–59%) in the maximum response was observed for twelve out of eighteen cells (67%) while three cells showed a mean 29% *increase* in the maximum response during cooling. The remaining three cells showed no change (<5%) in their maximum responses during cooling of A21a. This again indicates a primarily excitatory influence of feedback projections from A21a onto A17 cells ($P = 0.04$, $n = 18$). It was also apparent that the CRFs of some A17 cells were shifted considerably along the contrast axis, indicating a change in

the contrast sensitivity during inactivation of A21a. Of eighteen cells, nine showed a *decrease*, i.e. leftward-shift of the $C_{50} \geq 20\%$ (range 26–81%) and seven cells showed an *increase*, i.e. rightward-shift of the $C_{50} \geq 20\%$ (range 47–490%). Two cells were not affected, i.e. with a change less than 5%. There was no statistically significant difference between the C_{50} before and during cooling ($P = 0.68$, $n = 18$). Furthermore, the change in the maximum responses and that in the C_{50} was not correlated ($r = -0.03$, $P = 0.9$, $n = 18$). It was noted that in some cells during inactivation, the period of the rising phase or the dynamic range of the CRFs was also affected (Fig. 1f, g). However, when the dynamic contrast ranges were measured on a log scale between the contrasts that gave 20% (C_{20}) and 80% (C_{80}) of the maximal response, we did not find a significant difference ($P = 0.64$, $n = 18$) between the dynamic range before (1.22 ± 0.13 log unit) and during cooling of A21a (1.06 ± 0.12 log unit).

In the present study we compared also responses at low contrasts, C_{10} and C_{20}, which elicited 10 and 20% of the maximum responses respectively (Fig. 2a, filled and open diamonds), before and during cooling of A21a. During cooling the median response of 8.5 spikes/s before cooling was slightly reduced to 7.2 spikes/s ($P = 0.71$, $n = 36$). Thus, unlike the changes during attention [15] or concurrent stimulation of collinear flanks [16], the feedback from A21a did not systematically affect neuron' responses to low contrasts. This is probably due to the occurrence of both leftwards- and rightwards-shift of CRFs along the contrast axis during the inactivation. However, as shown earlier the magnitude of maximum responses was substantially reduced. Likewise, the responses at high contrasts, C_{80} and C_{90}, at which 80 and 90% of the maximum responses were elicited respectively (Fig. 2b,

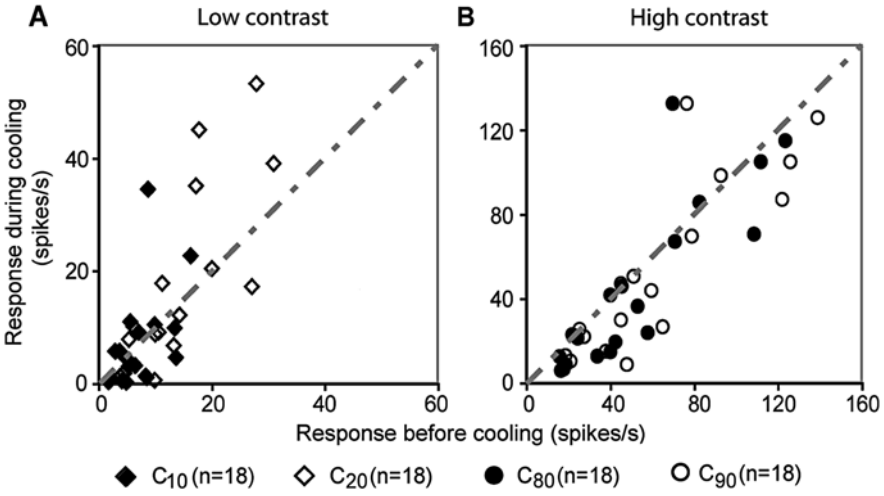


Fig. 2 Responses examined at both low (**a**) and high (**b**) contrasts before and during cooling. (**a**) Responses at low contrasts (C_{10} and C_{20}) were affected by inactivation of A21a. However, the change was not consistent for the 18 cells tested. (**b**) There was a systematic reduction in the responses at high contrasts (C_{80} and C_{90}) during inactivation.

filled and open circles), showed a significant difference between the responses before (57.2 ± 8.4; median: 46.4 spikes/s) and during (49.2 ± 9.5; median: 33.4 spikes/s) inactivation of A21a ($P < 0.006$, $n = 36$). Thus, the excitatory influence from A21a on CRFs of A17 cells was primarily at the top-end of the dynamic range. The increase in the magnitude of the maximum response without an expansion of dynamic range in stimulus contrasts would lead to a steeper slope on the rising part of CRF, that is, an increase in the magnitude of the responses per contrast unit. This, in turn, suggests an enhanced sensitivity to stimulus contrasts. Except for three cells showing a substantial change of more than two-fold in the slope (one decreased and two increased) we did not find a significant change in the slope during inactivation of A21a (*control*, median: 47.9 spikes/s/log unit of contrast; *cool*, median: 45.3 spikes/s/log unit of contrast; $P = 0.27$, $n = 18$). Lack of a noticeable change in the slope was probably due to changes (mostly a reduction) in the dynamic range observed during inactivation of A21a.

As reported here, in a substantial proportion (72%, 13/18) of A17 cells their CRFs were significantly affected during inactivation of A21a. The main effect was a reduction of the magnitude of responses at high contrast with a left- or right-ward shift of the CRFs and in some instances a narrowing in the contrast dynamic range, suggesting that the feedback from A21a influences the gain control mechanisms for both response and contrast.

Previous studies on the CRFs of A17 cells suggested that contrast responses can be modulated by simultaneously stimulating the classical RF and the surround region [17, 18]. The effect was generally a reduction in the magnitude of response. As suggested by Cavanaugh and colleagues [18], part of the surround suppression observed in their study may come from higher-order areas. In the present study, however, although we used long bars extending well into the surround region only two cells (2/18) showed an increase in the magnitude of maximum responses, implying a normally suppressive influence from A21a. This result is consistent with the findings in our previous study [5] where only a small proportion of end-stopped (hypercomplex) cells (2/10) became less suppressive during inactivation of A21a. Thus, the prevalent reduction in responses observed in the present study is more likely to be due to reduced excitatory inputs to these cells from A21a. On the other hand, a weak suppressive influence may also be present but masked by a predominantly excitatory feedback input.

Cortical cells are capable of adjusting contrast gain following adaptation to stimulus contrast, a mechanism which allows them to code an extremely wide range of environmental contrasts into a limited frequency range of neuronal discharges. A previous study comparing the monocular and binocular mechanisms of cortical contrast gain control, suggested that contrast adaptation occurs primarily before convergence of inputs from the two eyes [19]. More importantly, Solomon and colleagues [13] found direct evidence of contrast adaptation in virtually all magnocellular cells of the macaque dorsal lateral geniculate nucleus (LGNd; see also [12] in both X and Y cells of cat's LGNd), that, in turn, rely on the contrast adaptation of their retinal inputs [11]. While there is no clear consensus concerning the contribution made at each stage of visual information processing to contrast

adaptation, it appears that some of the effect observed in the cortex is generated intrinsically within cortical neurons [14, 20]. A profound change in contrast gain for some of our cells during inactivation of A21a (e.g. Fig. 1c, f) indicates that contrast gain control in A17 can be further modified by feedback projections. Presumably, due to a prevailing excitatory influence of A21a, we found that a change in contrast gain was very often associated with a change in response gain. Our current finding of a reduction in the dynamic range of CRF during inactivation of A21a (e.g. Fig. 1c, e–g) suggests an influence of A21a on the effective range of contrast detection of A17 cells. An extended dynamic range with increased neuronal responses would enhance the consistency of contrast discrimination over a wider range of contrasts. However, as indicated in Fig. 1f this enhancement was somehow achieved at the expense of detection at low contrasts. It is not clear how the visual cortex regulates the competing needs for detection sensitivity at low contrasts and contrast discrimination over a wide range of contrasts. It is, nevertheless, important to note that an excitatory influence of feedback projections on the magnitude of response could improve the quality, i.e. "signal-to-noise ratio" of the responses of A17 neurons wherever an increase in the magnitude of response is not at the expense of the deterioration in the relationship between the response and its variance [6].

References

1. Salin, P.A., Bullier, J.: Corticocortical connections in the visual system: Structure and function. Physiol. Rev. **75** (1995) 107–154.
2. Dreher, B.: Thalamocortical and corticocortical interconnections in the cat visual system: Relation to the mechanisms of information processing. In: Pettigrew, J.D., Sanderson, K., Levick, W.R., eds.: Visual Neuroscience. Cambridge: Cambridge University Press (1986) 290–314.
3. Mignard, M., Malpeli, J.G.: Paths of information flow through visual cortex. Science **251** (1991) 1249–1251.
4. Hupé, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., Bullier, J. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature **394** (1998) 784–787.
5. Wang, C., Waleszczyk, W.J., Burke, W., Dreher, B.: Modulatory influence of feedback projections from area 21a on neuronal activities in striate cortex of the cat. Cereb. Cortex **10** (2000) 1217–1232.
6. Wang, C., Waleszczyk, W. J., Burke, W., Dreher, B.: Feedback signals from cat's area 21a enhance orientation selectivity of area 17 neurons. Exp. Brain Res. **182** (2007) 479–490.
7. Huang, J.Y., Wang, C., Dreher, B.: The effects of reversible inactivation of postero-temporal visual cortex on neuronal activities in cat's area 17. Brain Res. **1138** (2007) 111–128.
8. Bardy, C., Huang, J.Y., Wang, C., FitzGibbon, T., Dreher, B.: 'Simplification' of responses of complex cells in cat striate cortex: suppressive surrounds and 'feedback' inactivation. J. Physiol. (Lond.) **574** (2006) 731–750.
9. Galuske, R.A., Schmidt, K.E., Goebel, R., Lomber, S.G., Payne, B.R.: The role of feedback in shaping neural representations in cat visual cortex. Proc. Natl. Acad. Sci. USA 99 (2002) 17083–17088.
10. Huang, L., Chen, X., Shou, T.: Spatial frequency-dependent feedback of visual cortical area 21a modulating functional orientation column maps in areas 17 and 18 of the cat. Brain Res. **998** (2004) 194–201.

11. Chander, D., Chichilinsky, E.J.: Adaptation to temporal contrast in primate and salamander retina. J. Neurosci. **21** (2001) 9904–9916.
12. Shou, T., Li, X., Zhou, Y., Hu, B.: Adaptation of visually evoked responses of relay cells in the dorsal lateral geniculate nucleus of the cat following prolonged exposure to drifting gratings. Vis. Neurosci. **13** (1996) 605–613.
13. Solomon, S.G., Peirce, J.W., Dhruv, N.T., Lennie, P.: Profound contrast adaptation early in the visual pathway. Neuron **42** (2004) 155–162.
14. Crowder, N.A., Price, N.S., Hietanen, M.A., Dreher, B., Clifford, C.W.G., Ibbotson, M.R.: Relationship between contrast adaptation and orientation tuning in V1 and V2 of cat visual cortex. J. Neurophysiol. **95** (2006) 271–283.
15. Reynolds, J.H., Pasternak, T., Desimone, R.: Attention increases sensitivity of V4 neurons. Neuron **26** (2000) 703–714.
16. Polat, U., Mizobe, K., Pettet, M.W., Kasamatsu, T., Norcia, A.M.: Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature 391 (1998) 580–584.
17. Sengpiel, F., Baddeley, R.J., Freeman, T.C., Harrad, R., Blakemore, C.: Different mechanisms underlie three inhibitory phenomena in cat area 17. Vision Res. **38** (1998) 2067–2080.
18. Cavanaugh, J.R., Bair, W., Movshon, J.A.: Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. J. Neurophysiol. **88** (2002) 2530–2546.
19. Truchard, A.M., Ohzawa, I., Freeman, R.D.: Contrast gain control in the visual cortex: monocular versus binocular mechanisms. J. Neurosci. **20** (2000) 3017–3032.
20. Sanchez-Vives, M.V., Nowak, L.G., McCormick, D.A.: Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. J. Neurosci. **20** (2000) 4267–4285.

Compensating Neural Transmission Delays by Dynamical Routing

Si Wu and Libo Ma

Abstract Neural delays are pervasive in the computations associated with brain functions. For adaptive behaviors, these delays must be compensated. The present study investigates a dynamical routing scheme which compensates the transmission delay of neural signals from retina to visual cortex. The scheme utilizes prior information of a moving stimulus and predicts the future position the stimulus will travel to after the delay period. We find that over a large range of speed, the routing scheme can compensate neural delays successfully.

Keywords Delay compensation · Dynamical routing · Neural computation · Visual information processing

1 Introduction

Neural delays are a general property of computations carried out by neural circuits. Given the visual system as an example, an external stimulus in the form of light intensity is first encoded into electrical signal by photo receptors in retina. This signal is further converted into action potentials by retinal ganglion cells. Subsequently, the neural signal is transmitted to LGN and finally reaches the primary visual cortex, where visual awareness of the stimulus is established. This whole process takes about 60–80 ms, which is a significant amount of delay in real-time computation. A fast moving object can travel a significant distance over the retinal surface during this period, implying that if this delay is not compensated properly, our perception of the location of a fast moving object will trail largely behind its true position. Thus, neural systems must have mechanisms to compensate neural delays.

S. Wu (✉)
State Key Laboratory of Neurobiology, Chinese Academy of Sciences, Institute of Neuroscience, Shanghai Institutes of Biological Sciences, Shanghai 200031, China
e-mail: siwu@ion.ac.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_68,

A possible strategy for a neural system achieving compensation is to predict the future position of a stimulus based on the prior knowledge it has on the motion speed of the stimulus. In a recent study, Nijhawan and Wu proposed a dynamical routing scheme which implements the compensation of the visual transmission delay [1]. Their idea is as follows: For a static stimulus, neural signal is transmitted along a straight line from retina to V1. Since V1 has the retinotopic map, the spatial position of the stimulus is properly registered. For a moving stimulus, on the other hand, neural signal is sent along a diagonal pathway to the visual cortex. The slope of the pathway is determined by the motion speed of the stimulus, so that the cortical position the neural signal is sent to corresponds to the future location the stimulus will travel to after the delay period. Hence, the transmission delay is compensated. An evidence supporting the dynamical routing scheme is the flash-lag effect observed in human psychophysical experiments [2].

Do the neural system has computational resources to achieve dynamical routing? Nijhawan and Wu found that this is possible [1]. They showed that at the regime of low speed, the asymmetric inhibitory interactions from amacrine starburst cells to ganglion cells is sufficient to implement dynamical routing. In this paper, we further study the dynamical routing scheme in the regime of high speed. In this case, the effect of asymmetrical inhibition from amacrine cells becomes negligible. Rather, it is the property that the time constant of inhibitory inputs is larger than that of excitatory ones contributes to routing behaviors. We demonstrate that over a large range of speed values, the dynamical routing scheme can compensate the visual transmission delay successfully.

2 Dynamical Routing in Visual Information Processing

We consider a neural circuit formed by retinal ganglion cells. The real network is two dimensional. Here, for simplicity, we consider a stimulus is moving in a straight line and only study a one-dimensional network.

Let us denote $r(x, t)$ the activity of ganglion cells at the retinal position x and at the time t. The ganglion cells receive excitatory inputs from bipolar cells and inhibitory ones from amacrine cells, which are denoted as $I_b(x, t)$ and $I_a(x, t)$, respectively. The activity of ganglion cells is determined by

$$r(x,t) = \frac{I_b(x,t)}{\rho + I_a(x,t)}, \tag{1}$$

where ρ is a small positive number, and the division operation represents the effect of multiplicative inhibition from amacrine cells to ganglion cells.

The stimulus position is encoded in the population activity of ganglion cells [3]. We use Center of Mass strategy to read-out this information, which is given by the center of $\{r(x, t)\}$, for all x values, i.e.,

$$C_r = \frac{\int_{-\infty}^{\infty} xr(x,0)dx}{\int_{-\infty}^{\infty} r(x,0)dx}. \tag{2}$$

The inputs from bipolar cells to ganglion cells are given by

$$I_b(x,t) = \int_{-\infty}^{\infty} dx' \int_{-\infty}^{t} dt' K_b(x-x', t-t') S_b(x', t'), \tag{3}$$

where $S_b(x', t')$ is the responses of bipolar cells to a moving stimulus. $K_b(x-x', t-t')$ is the receptive field of ganglion cells with respect to bipolar cells,

$$K_b(x-x', t-t') = h_b e^{-\frac{(x-x')^2}{2a_b^2}} e^{-\frac{t-t'}{\tau_b}}, \tag{4}$$

where h_b is a constant, a_b the width of receptive field, and τ_b the time constant of excitatory inputs.

The inputs from amacrine cells to ganglion cells are

$$I_a(x,t) = \int_{-\infty}^{\infty} dx' \int_{-\infty}^{t} dt' K_a(x-x', t-t') S_a(x', t'), \tag{5}$$

where

$$K_a(x-x', t-t') = h_a e^{-\frac{(x-x')^2}{2a_a^2}} e^{-\frac{t-t'}{\tau_a}}, \tag{6}$$

and all variables are similarly defined as in Eq. (4).

We consider a stimulus is moving at a constant speed v, and its position at time t is z. Hence, $x(t') = z - v(t-t')$.

In the previous study [1], the authors considered the situation when the speed of the stimulus is sufficiently slow, in term of that $v\tau_a$ is in the order of the size of a single amacrine cell. In this case, the asymmetric inhibitions from amacrine cells to ganglion cells contribute to dynamical routing. Here, we are interested in the regime when the speed of the stimulus is sufficiently large, so that the stimulus crosses a single amacrine cell very quickly, and ganglion cells are unable to differentiate asymmetric inhibitory interactions. For the convenience of analysis, we consider the instant responses of amacrine and bipolar cells to the stimulus are the same, i.e.,

$$S_b(x', t') = S_a(x', t') = \delta(x' - z + v(t-t')), \tag{7}$$

where the symbol $\delta(x)$ denotes the delta function.

Without loss of generality, we calculate the population activity of ganglion cells at time $t = 0$ and assume $z = 0$ at $t = 0$. We get

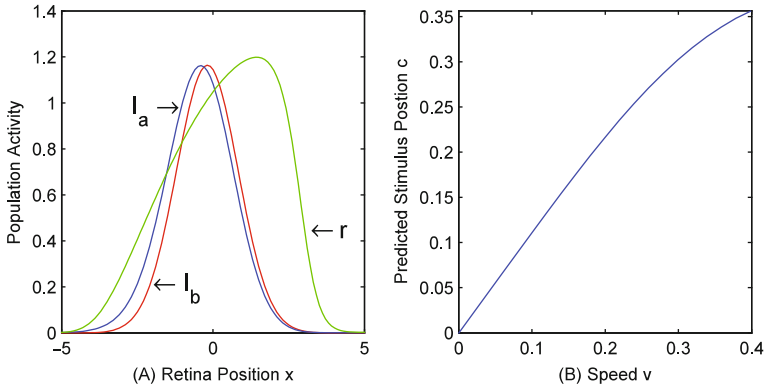


Fig. 1 (**a**) The population activities of ganglion, bipolar and amacrine cells; (**b**) The predicted stimulus position vs. the motion speed of the stimulus

$$I_b(x,0) = \int_0^\infty h_b e^{-\frac{(x+vt)^2}{2a_b^2}} e^{-\frac{t}{\tau_b}} dt, \tag{8}$$

$$I_a(x,0) = \int_0^\infty h_a e^{-\frac{(x+vt)^2}{2a_a^2}} e^{-\frac{t}{\tau_a}} dt. \tag{9}$$

As shown in Fig. 1a, both $\{I_b(x,0)\}$ and $\{I_a(x,0)\}$ have a bump shape tilted to the opposite direction of the stimulus moving. Their centers are given by $C_b = -\tau_b v$ and $C_a = -\tau_a v$, respectively. From Eq. (2), we see that if the inhibitory bump is tilted more towards the opposite direction of stimulus moving than the excitatory bump, then the bump of ganglion cell activities is tilted towards the moving direction of the stimulus (Fig. 1a), implying that the network predicts the future position of the stimulus. A proper compensation requires that the center of the bump $\{r(x,0)\}$, i.e., C_r, is proportional to the speed of the stimulus.

Let us consider an analytically solvable case when $\tau_1 v << a_1$ and $\tau_2 v << a_2$. Under this condition, the distortions of the bumps I_b and I_a from a Gaussian function are small, we may write down

$$I_b(x,0) \approx h_b \tau_b e^{-x^2/2a_b^2}(1 - \frac{\tau_b v x}{a_b^2}), \tag{10}$$

$$I_a(x,0) \approx h_a \tau_a e^{-x^2/2a_a^2}(1 - \frac{\tau_a v x}{a_a^2}). \tag{11}$$

Thus,

$$r(x,0) \approx \frac{h_b \tau_b}{h_a \tau_a} e^{-x^2/2a_b^2 + x^2/2a_a^2}(1 + \frac{\tau_a a_b^2 - \tau_b a_a^2}{a_a^2 a_b^2} v x), \tag{12}$$

whose center is calculated to be $C_r = (\tau_a a_b^2 - \tau_b a_a^2)/(a_a^2 - a_b^2)v$. Thus, we see that if $a_a > a_b$ and $\tau_a a_b^2 > \tau_b a_a^2$, i.e., the time constant of inhibitory inputs is sufficiently larger than that of excitatory ones, C_r is proportional to v.

Figure 1b shows that with proper parameters, the stimulus position predicted by the network increases linearly with the speed of the stimulus over a large range of speed values, implying the compensation of the delay is achieved.

3 Conclusions

To conclude, we have investigated a dynamical routing scheme for compensating the visual transmission delay. The scheme utilizes prior information of a moving stimulus and dynamically selects a route to send neural signal from retina to the visual cortex. The scheme works well over a large range of speed values.

References

1. Nijhawan, R., Wu, S.: Philos. Trans. R. Soc. A **367** (2009) 1063–1078.
2. Nijhawan, R.: Nature **370** (1994) 256–257.
3. Berry, M. et al.: Nature **398** (1999) 334–338.

Retinotopic Sparse Representation of Natural Images

Libo Ma

Abstract Independent component analysis and sparse coding have provided a functional explanations of simple cells in primary visual cortex (V1). The learned components(corresponding to the responses of neurons) of these models are randomly scattered and have no particular order. In practice, however, the arrangement of neurons in V1 are ordered in a very specific manner. In this paper, we propose a sparse coding of natural images under a retinotopic map constraint. We investigate the spatial specifically connections between retinal input and v1 neurons. Some simulations on natural images demonstrate that the proposed model can learn a retinotopic sparse representation efficiently.

1 Introduction

Efficient coding hypothesis has been considered as a general computational criteria for the modeling of early visual system. Over the last decade, a variety of methods approach the computational modeling of primary visual cortex based on statistical modeling of natural images. Independent component analysis [1] and sparse coding [2] have shown that the statistically optimal linear features of natural images are very similar to the receptive fields (RFs) of simple cells in primary visual cortex [2–5].

Algorithms for learning efficient coding commonly divide natural images into small patches as the input because the computational costs become heavier for large scale images. And the learned basis functions or coefficients have no particular order, or other relationships. This lack of an intrinsic order of basis functions is correlated with the assumption of complete statistical independence.

L. Ma (✉)
State Key Laboratory of Neurobiology, Chinese Academy of Sciences, Institute of Neuroscience, Shanghai Institutes of Biological Sciences, Shanghai 200031, China
e-mail: malibo@ion.ac.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_69,

However, in the most visual areas of brain, especially in the primary visual cortex, there is a prevalent feature that the visual world is mapped onto the cortical surface in a topographic manner. The neurons in primary visual cortex are arranged orderly according to the retinotopic map. This retinotopic map refers to the transformation from the coordinates of the visual world to the corresponding locations on the cortical surface. Namely, neighboring points in a retinal visual image arouse activity in adjacent neurons of primary visual cortex. Furthermore, neighboring neurons have RFs that include overlapping portions of the visual field. The position of the center of these RFs forms an orderly sampling mosaic that covers a part of the visual field. Thus, it is important in the modeling of primary visual cortex that the retinotopy is utilized during estimating the basis functions(corresponding to receptive fields), so that the estimated components can be ordered in a meaningful way.

In this paper, we propose a modification of sparse coding that explicitly formalizes a retinotopic map representation on the basis functions. This gives a retinotopic map where the position of the receptive field of V1 neurons emerge from the spatial specificity of connections between retinal input images and V1 neurons. This gives a new principle for retinotopic organization. Moreover, we derive a new learning algorithm for the evaluation of the method. We demonstrate that over a comparatively large natural images, the retinotopic map representation are learned successfully.

2 Retinotopic Sparse Representation Method

Now we define a generative model that implies retinotopic map for visual neurons that have local receptive field and overlapping between adjacent neurons. A simple form of coding strategy is a linear factor model, in which a static monochrome image $I(x, y)$ is assumed to be generated as a linear superposition of basis functions $a_i(x, y)$ with additive noise $\varepsilon(x, y)$ as:

$$I(x, y) = \sum_{i=1}^{n} a_i(x, y)s_i + \varepsilon(x, y) \tag{1}$$

where s_i are neuronal activities when presented with a given image I. This is a classical sparse coding [2].

In order to formalize a retinotopic representation in the model, we apply a retinotopic map function $C_{i,j}(x, y)$ to be multiplied on each basis function as:

$$I(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}(x, y)a_{i,j}(x, y)s_{i,j} + \varepsilon(x, y) \tag{2}$$

A piece of cortex is modeled by a two-dimensional grid of $N \times N$ neural units ("neurons"). The units should be interpreted as populations of neighboring cortical neurons, not as a single neuron.

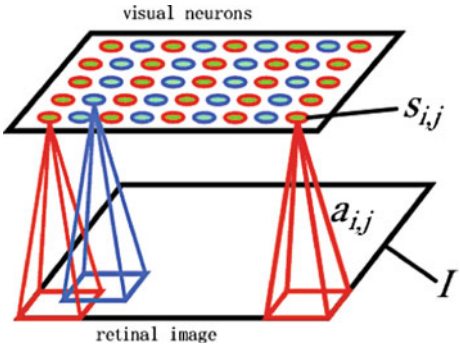


Fig. 1 An illustration of retinotopic representation model network

Let the size of a given natural image I is $M \times M$. We define a retinotopic map function, $C_{i,j}(x, y)$, as a projection constraint on the connections from retinal image of position (x, y) to the (i, j) visual neuron. The center of receptive field of the (i, j) neuron is located at (x, y) and its size is $m \times m$, and $m << M$. This means that the (i, j) neuron can only receive small local field from input retinal image with specific position and size. Moreover, we assign adjacent neurons have receptive fields that include overlapping portions of the input retinal image field. That is, the neighboring patch size of $m \times m$ at position $(x, y+1)$ are projected onto neighboring neuron $(i, j+1)$ correspondingly. As a result, the centering position of these receptive fields form an orderly sampling mosaic that covers the whole input retinal image. See Fig. 1. Furthermore, all the $N \times N$ neighboring patches which surround the center (x, y) are projected onto $N \times N$ neurons in succession. So, we can obtain $M = N + m - 1$.

The estimation of the model is similar to classical sparse coding, see [2]. Now, in this model, the objective function can be written as:

$$L = \sum_{x,y} [I(x, y) - \sum_{i,j} C_{i,j}(x, y) a_{i,j}(x, y) s_{i,j}]^2 + \lambda \sum_{i,j} S(s_{i,j}) \tag{3}$$

where λ is a constant. The minimization of objective function L with respect to $a_{i,j}$ and $s_{i,j}$ can be divided into two stages. First, to estimate basis functions $a_{i,j}$ to find a retinotopic representation of natural images with fixed $s_{i,j}$. Second, to infer the coefficients $s_{i,j}$ for a given image I, holding the basis functions $a_{i,j}$ fixed. Thus, by using gadient descent algorithm, the learning rule for basis functions and coefficients can be accomplished. Note that, we set the norm of basis function to be unity after each updating on the basis functions.

3 Simulations

In this section, we test retinotopic sparse representation on natural images that is distributed by Olshausen [2]. The size of receptive field of each neuron is set to be

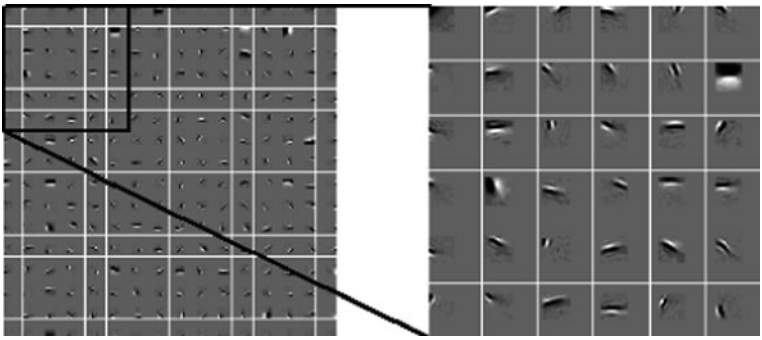


Fig. 2 Part of synaptic weights W by local independent factorization algorithm. The enlarged synaptic weights that is shown by *dashed square*

$m = 16$. The number of neurons denoted as $N = 16$. So the size of input image is $M = 31$. These images are prewhitened to equalize large variances in all directions in frequency domain. We apply batch mode to train the model, in which three hundred image patches are extracted in each batch.

The learned basis functions or receptive fields by retinotopic sparse representation are shown in Fig. 2. The local region shown by solid square is enlarged to show the details of basis functions. We can see that the proposed model can learn Gabor-like classical receptive fields (CRFs) of simple cells. A distinct feature of the model is that the receptive field of neuron is enclosed in a local area around the center of neuron approximately.

4 Conclusions

We have introduced a new sparse coding method that we embed a retinotopic map on the basis function to model the mapping relationship between retinal image and visual neurons. We derived a Learning algorithm for training the model by using gradient descent algorithm. Simulations demonstrate that the learned basis functions are similar to the Gabor-like classical receptive fields of simple cells.

References

1. Bell, A.J., Sejnowski, T.J.: The 'independent components' of natural scenes are edge filters. Vis. Res. **37**(23) (1997) 3327–3338.
2. Olshausen, B., Field, D.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**(6583) (1996) 607–609.
3. van Hateren, J.H.: Independent component filters of natural images compared with simple cells in primary visual cortex. Proc. Biol. Sci. **265**(1394) (1998) 359–366.

4. van Hateren, J.H.: Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proc. R. Soc. B Biol. Sci. **265**(1412) (1998) 2315–2315.
5. Caywood, M., Willmore, B., Tolhurst, D.: Independent components of color natural scenes resemble V1 neurons in their spatial and color tuning. J. Neurophysiol. **91**(6) (2004) 2859–2873.

A Neural Network Model for Image Formation of the Face Through Double-Touch Without Visual Information

Wei Song, Toshio Inui, and Naohiro Takemura

Abstract The fact that neonates only 9 min old preferentially look at faces suggests that they have already acquired the visual schema of a face. However, the mechanism for learning the schema of "face" during the fetal period is still unknown. In this paper, a neural network model that can acquire visual schema without visual input is proposed, which is based on the fact that fetuses frequently touch their own faces. This model consists of three major functions: (1) transmission of topological information about touched area on the face to visual area through VIP, which is realized by a 2-layered SOM with Gaussian-like local connections, (2) coding of shape information with Gabor filters by finger-touching, and (3) integration of two kinds of information (i.e., topological and shape information) using Hebbian connection. The sum of Hebbian activation appears to be similar to the representation of face in polar-coordinates.

Keywords Neonate face preference · Neural network model · Double touch

1 Introduction

From various studies on infant facial imitation, it is apparent that neonates have the body scheme of their faces. Fuke and her colleagues suggested that neonates acquire their visual facial schema from ego-centric visual information of their hand-position [1, 2]. However, newborn infants of only 9 min age already have a preference for human facial patterns, suggesting that they acquire visual facial schema during their fetal period [3]. In this paper, we propose a new model for acquisition of a visual facial image without visual information.

W. Song (✉)
Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, 606-8501 Kyoto, Japan
e-mail: soui@cog.ist.i.kyoto-u.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_70,

2 Outline of the Model

Fetuses often touch their faces with their hands (double-touch). We assume that a visual facial schema is acquired through double-touch, and propose a model of facial image formation without vision. When a part of a fetus's face is touched by his/her hand, the tactile receptors on the facial skin are activated. Topological information about touched area on the face is transmitted to, and represented in, the VIP (ventral intraparietal area) via an onion-like map in the primary somatosensory cortex (SI) [4]. In this model, a somatotopic face map in the VIP, which was found in physiological experiments on the macaque monkey, is learned through face touching. Here, we assumed that there are topological connections from the VIP to the visual cortex, so that somatotopic information regarding the touched area on the face is transmitted to the visual cortex via the VIP. On the other hand, touching the face with his/her finger generates tactile information of surface properties of the touched facial area, and this is encoded by Gabor filters in the SI. The positional information about the touched area of the face and the shape information of the face from finger are integrated with a Hebbian learning algorithm. (see Fig. 1)

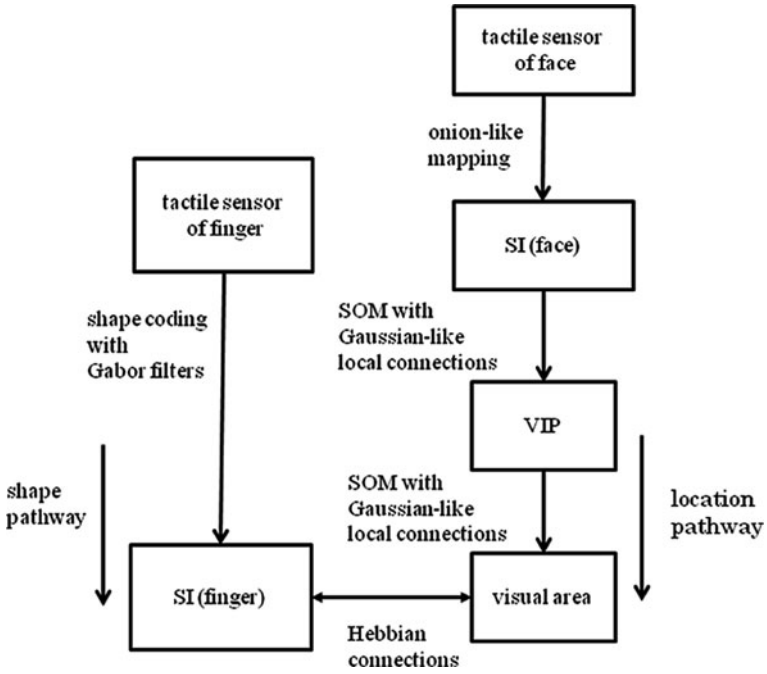


Fig. 1 Schematic diagram of the model

3 Detail of the Model

3.1 Location Pathway

When a certain part of the face is touched by the finger, tactile sensors within the area of the finger size are activated. Each tactile receptor projects to a corresponding SI unit. The locations of the units in the SI map represent onion-like topology of face (Fig. 2). In the model, connections between SI and VIP as well as VIP and visual cortex are learned with the self-organizing map (SOM). The structures of the SOM are assumed to be local Gaussian-like local connections (Fig. 3). This assumption is similar to the self-organizing model of the retino-cortical pathway [5].

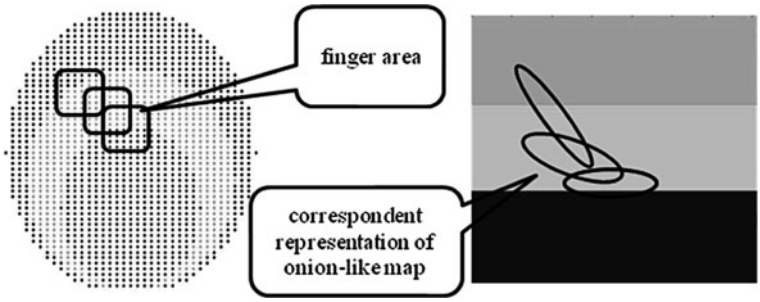


Fig. 2 Corresponding areas for finger on face and SI map

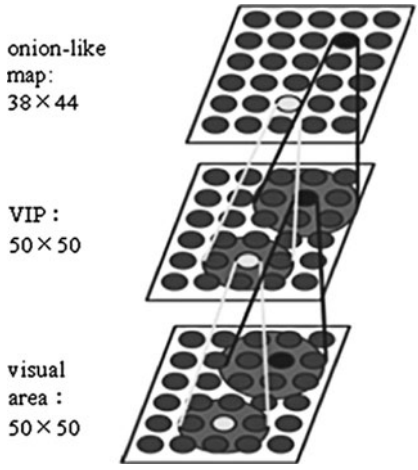


Fig. 3 Structure of 2-layered SOM

3.2 Shape Pathway

In this model, concavo-convex information of the face within the finger surface is encoded by Gabor filters in SI according to the physiological findings [6]. There are 16 kinds of filters with different orientation selectivity (Fig. 4). The population of finger SI units codes shape information for each location within this filter bank.

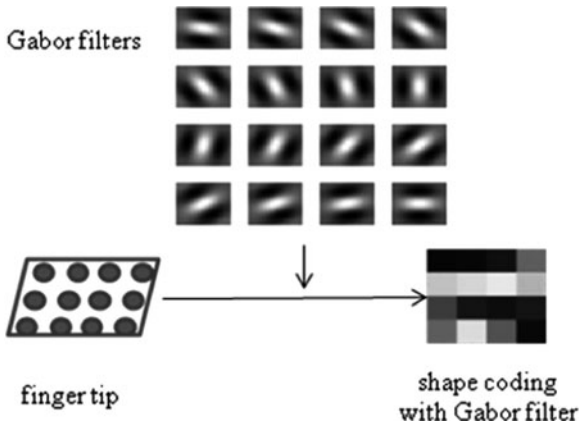


Fig. 4 Gabor filters and shape coding

3.3 Integration of Two Kinds of Information

When a certain area of the face is touched, the information for location and surface of the touched area of the face are encoded in the location pathway and the shape pathway, respectively. As shown in Fig. 1, these two types of information are integrated by a Hebbian learning algorithm between the shape pathway and the visual cortex.

4 Results and Discussion

After unsupervised learning of the SOM, VIP units acquired facial somatotopic mapping (Fig. 5) similar to that seen in human fMRI data [7]. Figure 6 shows the acquired facial image in the visual cortex. This image appears when all the SI (finger) units are activated. In other words, it is a potential image formed by a number of finger tactile inputs.

Retino-cortical mapping is well known as a log-polar transformation of the retinal image (e.g., Shwartz [8]; Fig. 6 right). It appears that the acquired facial image is similar to the visual representation in a log-polar transformation of the face. The visual facial schema can be acquired in the fetal period through double-touch

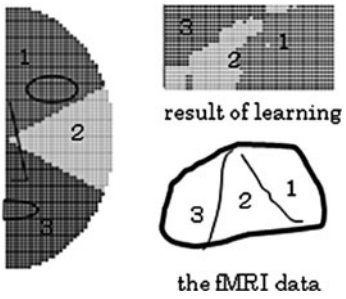


Fig. 5 Comparison between learning result and the fMRI data [7] of VIP map

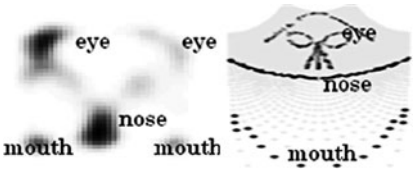


Fig. 6 Learning result of Hebb and the face representation in polar coordinates

between the finger and face. This schema may be the basis for the preferential looking for faces and for facial imitation.

References

1. Fuke, S., Ogino, M., Asada, M. Body image constructed from motor and tactile images with visual information. Int. J. Humanoid Robotics **4** (2007) 347–364.
2. Fuke, S., Ogino, M., Asada, M. Visuo-tactile face representation through self-induced motor activity. Proceedings of the 7th International Conference on Epigenetic Robotics.
3. Goren, C.C., Sarty, M., Wu, P.Y.K. Visual following and pattern discrimination of face-like stimuli by newborn infants. Pediatrics **56** (1975) 544–549.
4. Moulton, E.A., Pendse, G., Morris, S., Aiello-Lammens, M., Becerra, L., Borsook, D.: Segmentally arranged somatotopy within the face representation of human primary somatosensory cortex. Hum. Brain Map. **30** (2009) 757–765.
5. Linsker, R.: Local synaptic learning rules suffice to maximize mutual information in a linear network. Neural Comput. **4** (1992) 691–702.
6. Bensmaia, S.J., Denchev, P.V., Dammann III, J.F., Craig, J.C., Hisiao, S.S. The representation of stimulus orientation in the early stages of somatosensory processing. J. Neurosci. **28** (2008) 776–786.
7. Sereno, M.I., Huang, R.S. A human parietal face area contains aligned head-centered visual and tactile maps. Nat. Neurosci. **9** (2006) 1337–1343.
8. Schwartz, E.L. Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. Biol. Cybern. **25** (1977) 181–194.

Population Responses of Aδ-Fiber Nociceptors with Hypothetical Spatial Distributions on Human Forearm

Jiyong Hu, Yi Li, Xin Ding, and Rubin Wang

Abstract This paper studied the effect of innervation patterns of cutaneous mechanosensitive Aδ-nociceptors with myelinated axons on the evoked neural responses while the skin was mechanically stimulated using quantitative indentation (compression). By reconstructing the neural response of nociceptors under a punctate stimulus, the peak firing rate and the number of recruited nociceptors are calculated for different innervation patterns. And then, these two measures are used to discuss the effect of external stimulus features on the neuronal activity. The results demonstrate that there are large differences between the active level of cutaneous Aδ nociceptors with Gaussian and other distribution patterns (i.e. rectangular and uniform). When the innervation density distribution is Gaussian, the stimulus intensity and contactor size can be well encoded by the proposed two measures.

Keywords Nociceptors · Innervation · Distribution · Sensitivity · Forearm

1 Introduction

Cutaneous nociceptors are one of the common tactile receptors. It is traditionally believed that the cutaneous myelinated Aδ-fiber nociceptors are responsible for fabric-evoked prickle, and the mean firing rates are the principle response feature that the central nervous system (CNS) uses to interpret a mechanical stimulus as prickling and painful [1]. Meanwhile, Garell et al. shown that the psychophysical magnitude estimation of mechanical stimulus intensity is consistent with population responses of nociceptors, not those of single Aδ-fiber nociceptors. Population activities depend on the distribution density and the mechanical sensitivity of nociceptors [1–3], and on the stimulus intensity and the contactor size which delivers the

J. Hu (✉)
Key Lab of Textile Science and Technology, Ministry of Education, Donghua University, Shanghai 201620, P.R. China
e-mail: hjydhu@126.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_71,

stimulus to human skin. In terms of the limitation of the present multi-unit recording techniques, these parameters have rarely been analyzed in detail and also systematic measurements have been not made to examine how they affect neuronal population responses.

The alternative method is to reconstruct the neural response of populations of nociceptors. Khalsa et al. represented the uniform population responses of nociceptors with the same threshold and the same mechanosensitivity in the rat hind thigh [2]. However, anatomical data are showing that Aδ-nociceptors can be randomly [3] or normally [4] distributed in epidermis. Furthermore, the assumption of the same threshold and the same sensitivity is not consistent with physiological property of nociceptors [1, 3]. Considering these factors, in this paper, the effect of innervation patterns of cutaneous Aδ nociceptors on the neural activity will be discussed by a punctate probe delivering the indentation vertical to human forearm skin.

2 Methods

The reconstruction is made by using Güçlü's technique [5], and the following relationships describing the response of single nociceptive endings are defined:

(1) A punctate stimulus deforms the extensive skin including contact sites. The effective deformation (D) under a flat-tip stylus is defined [6] as

$$D(r) = \begin{cases} \dfrac{p}{2r_f}\dfrac{1-v^2}{E} & r \le r_f \\ \dfrac{p}{\pi r_f}\dfrac{1-v^2}{E}\sin^{-1}(r_f/r) & r > r_f \end{cases} \tag{1}$$

Where p and r_f is the force on skin by the probe and its radius; E and v is the Young modulus and Poisson's ratio of forearm skin.

(2) According to previous work [2–4], three hypothetical innervation distributions are used to represent the receptive-field organization: rectangular, uniformly random and Mid-lateral Gaussian. The corresponding x- and y-coordinates of the receptive field centers are randomly determined by using the classical independent probability density functions. For Gaussian random distribution:

$$f_X(x) = \frac{\alpha}{\sigma\sqrt{2\pi}} e^{-(x-x_c)^2/2\sigma^2} + \beta \tag{2}$$

and

$$f_X(x_c) = \frac{\alpha}{\sigma\sqrt{2\pi}} + \beta = d_{\max},\ \lim_{x\to\infty} f_X(x) = d_{\min},\ \frac{1}{w}\int_0^w f_X(x)dx = d \tag{3}$$

Where d-average density per mm^2, d_{min} and d_{max}-the minimum and maximum density, respectively, w-the width of forearm skin, σ is the standard deviation,

α and β are the two other parameters. Note that the x-axis refers to the medial-lateral, the y-axis to the distal-proximal axis. Meanwhile, individual Aδ-fibers branch into many terminals which innervate many nociceptive cells, and each fiber allows for a uniform random number of sensitive spots within a receptive field.

The statistical distribution density of cutaneous Aδ-nociceptors can just be estimated by the proportion of single type of afferents in human skin. The average SAI units in the human forearm are 4 fibers per cm^2 [7], which approximately represent 38% of the myelinated afferent population, probably with equal numbers of Aβ and Aδ fibers [6]. Twenty percent of all Aδ receptors are nociceptors and 50% of these are mechanically sensitive. Allowing for each Aδ fiber in 3–11 spots (say an average of 7.8 for a single Aδ fiber) [2], there are averaging 0.675 Aδ-nociceptive afferents and 5.265 nerve ending organs per cm^2. Thus, all the parameters in Eq. (3) can be solved in real-time with MATLAB during each simulation run.

(3) The stimulus (force F) – response (evoked firing rates, f) relationship of Aδ-fiber nociceptors is described by Khalsa et al. as

$$f(F) = a(1 - e^{-bF}) + c \tag{4}$$

Where a, b and c were parameters to be fit during nonlinear regression [1]. The result is a=–11.32, b=0.04 and c= 13.22.

3 Results and Discussion

The probability of recording one or more Aδ-fiber nociceptors at an individual recording site strongly depends on innervation density of Aδ-fiber endings. To quantitatively compare the effect of innervations pattern on the evoked neural responses and the recruited number of nociceptive units in the same stimulus condition, they are potted as the function of loading force in Fig. 1, respectively. Here, the activated fiber number under single mechanical stimulus is counted if its firing rate is

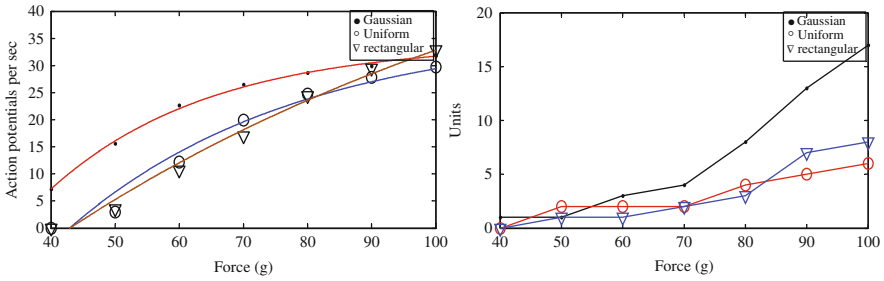


Fig. 1 Neural responses of nociceptors under different stimulus intensity and innervations patterns. *Left*: peak action potentials per second; *right*: Recruited numbers of nociceptive fibers

not less than 1.0 Hz. In Fig. 1, the absolute firing rate of single punctate stimulus is highest for the nociceptive units with Gaussian peak distribution, especially when the stimulus intensity is not much stronger than Von Frey filament threshold. With the increasing stimulus intensity less than pain threshold, the effect of innervation pattern on the evoked firing rates become minute, and the recruited numbers of nociceptive units increase monotonically, especially for Gaussian distribution. When the stimulus intensity is weak, there are no significant differences of the recruited numbers among three innervation patterns. In compare with the peak neural response of populations of nociceptive units, the more numbers of recruited nociceptive units do not change the peak response magnitude. This trend is consistent with the results by Garell et al. [1]. Greenspan observed that the perception of sharpness show no significant spatial summation and is commonly perceived by near mechanothreshold activity of single Aδ-fiber nociceptors.

Figure 2 shows that the peak action potentials of populations of nociceptive afferents increase with an increase of probe size, and the number of recruited nociceptive afferents increases monotonically. Especially, the activity of nociceptors in Gaussian peak distribution is dramatically different from that in either uniform or rectangular distribution. Meanwhile, the probe size less than 0.05 mm in Gaussian distribution, and less than 0.5 mm in either uniform or rectangular distribution has nearly no effect on the neural activity of nociceptors. Especially, the increasing peak response rather than the number of recruited afferents implies the spatial summation across single RFs for probe less than 0.5 mm. Garell et al. observed that the peak response is higher for probe with little radius than that with large radius [1].

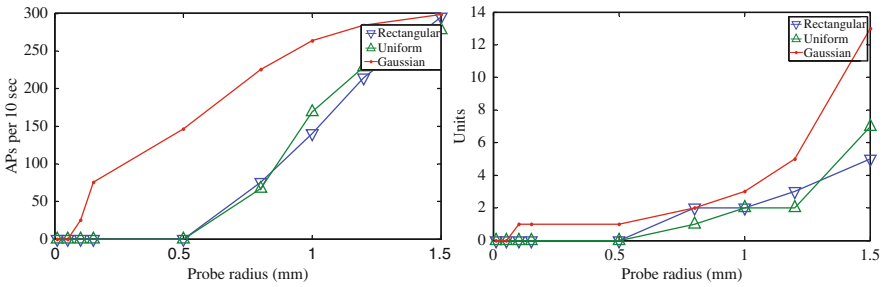


Fig. 2 Effect of probe radius on neuronal responses of populations of nociceptors with different innervation patterns. *Left*: peak action potentials; *right*: recruited numbers of nociceptive fibers

4 Conclusion

This study concluded that innervation pattern of nerve endings have significant effect on peak neural responses of populations of nociceptors. The recruited numbers of nociceptive units potentially encode the external stimulus intensity and

contactor size, and the location of peak neural response occurrence potentially encodes the stimulus site. Meanwhile, the probe radius delivering stimulus has significant effect on nociceptors with Gaussian innervation patterns in a limited range.

Acknowledgments The authors would like to acknowledge the project (No. KF0903) and NNSF through project (No.10672057 and No. 10872068).

References

1. Garell, P.C., McGillis, S.L.B., Greenspan, J.D.: Mechanical response properties of nociceptors innervating feline hairy skin. J. Neurophysiol. **75**(3) (1996) 1177–1189.
2. Khalsa, P.S., Zhang, C., Qin, Y.X.: Encoding of location and intensity of noxious indentation into rat skin by spatial populations of cutaneous mechano-nociceptors. J. Neurophysiol. **83** (2000) 3049–3061.
3. Slugg, R.M., Campbell, J.N., Meyer, R.A.: The population response of A- and C-Fiber nociceptors in monkey encodes high-intensity mechanical stimuli. J. Neurosci. **24**(19) (2004) 4649–4656.
4. McArthur, J.C., Stocks, E.A., Hauer, P., et al.: Epidermal nerve fiber density: Normative reference range and diagnostic efficiency. Arch. Neurol. **55**(12) (1998) 1513–1520.
5. Güçlü, B., Bolanowski, S.J.: Modeling population responses of rapidly-adapting mechanoreceptive fibers. J. Comput. Neurosci. **12** (2002) 201–218.
6. Johnson, K.L.: Contact Mechanics. Cambridge: Cambridge University Press (1985)..
7. Vallbo, Å.B., Olausson, H., Wessberg, J., et al.: Receptive field characteristics of tactile units with myelinated afferents in hairy skin of human. J. Physiol. **483**(3) (1995) 783–795.

Optimal Model for Selecting Human Arm Posture During Reaching Movement

Jiaole Wang and Masazumi Katayama

Abstract In this research, we investigated the optimization principles for human arm posture selection during the reaching movement in a horizontal plane. Because the human arm has infinite possible postures at a same hand position, we evaluated three optimization models for arm posture selection: the minimum angular jerk model, the minimum torque change model, and the minimum commanded torque change model. After quantitatively comparing measured arm postures with the optimal ones calculated by each optimization model, we found that the minimum commanded torque change model is more plausible for human arm posture selection.

Keywords Arm posture selection · Optimal model · Reaching movement

1 Introduction

When we are doing reaching movement by using our upper limbs, we are supposed to solve three kinds of computational problems: trajectory formation, coordinates transformation, and control [1]. Although all these three ill-posed problems have infinite possible solutions, humans show highly stereotyped motor behaviour and these common invariant spatiotemporal features of planar reaching movements have been reported by many researchers [2, 3].

For the trajectory formation problem of the human arm regarded as a 2-link manipulator, from the computational point of view, the hand trajectory should be planed based on an optimal principle so that only one hand trajectory is determined from the infinite ones. Many researchers have proposed and evaluated lots of principles such as the minimum hand jerk model [3], the minimum angular jerk model [4], the minimum torque change model [5], the minimum commanded torque change model [6], and the minimum variance model [7]. As a result, some researchers have reported that the minimum commanded torque change model and the minimum

M. Katayama (✉)
University of Fukui, Bunkyo 3-9-1, Fukui, Japan
e-mail: kata@u-fukui.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_72,

variance model are more plausible. Thus, the trajectory formation problem has been successfully elucidated.

While the problem of coordinates transformation for the human arm has not been investigated well, although in the robotics field it has been investigated adequately. As we all know, because the human arm can be seen as a redundant manipulator, there are infinite possible postures during a movement. However, we only do one posture at one hand position during our movements, there must be a principle for us to select one posture from infinite possible ones. Tsuyuki et al. [8] have investigated this problem for 2-link arm in the 3D space and reported that the minimum commanded torque change model is more plausible.

In this research, we focus on the coordinates transformation problem for the human arm regarded as a 3-link manipulator in the horizontal plane, that is an arm posture selection problem for the reaching movement. We respectively compare the measured arm posture with each optimal posture predicted by three optimization models which are thought to be plausible: the minimum angular jerk model, the minimum torque change model and the minimum commanded torque change model.

2 Optimization Models for Arm Posture Selection

In this research, we evaluated the below three optimization models that select the optimal arm posture by minimizing the value of each objective function.

Minimum angular jerk model (AJ): Rosenbaum et al. [4] proposed the minimum angular jerk model for solving the trajectory formation problem. The angular jerk is defined as differentiating the joint angle θ three times with respect to time t. θ_i stands for the angle of ith joint out of n joints, t_f is the movement duration.

$$C_{AJ} = \frac{1}{2}\int_0^{tf} \sum_{i=1}^{n} \left(\frac{d^3\theta_i}{dt^3}\right)^2 dt. \tag{1}$$

Minimum torque change model (TC): Because the minimum jerk model has only taken the kinematics into account, Uno et al. [5] proposed the following alternative optimization model: movements are organized such that the time integral of the squared sum of torque changes is minimal. τ_i stands for the joint torque of the ith joint.

$$C_{TC} = \frac{1}{2}\int_0^{tf} \sum_{i=1}^{n} \left(\frac{d\tau_i}{dt}\right)^2 dt. \tag{2}$$

Minimum commanded torque change model (CTC): By extending TC, Nakano et al. [6] proposed the minimum commanded torque change model which is expressed as below:

$$C_{CTC} = \frac{1}{2}\int_0^{tf} \sum_{i=1}^{n} \left(\frac{d\tau_{Ci}}{dt}\right)^2 dt, \tag{3}$$

τ_{Ci} is the commanded torque defined as $\tau_C = \tau + \dot{\theta} B(\tau)$ [6], by employing motor commands at the peripheral level. B stands for viscous coefficient matrix around joint which depends on the actual torque τ. Here, the diagonal components of B which are B_{11}, B_{22} and B_{33} are the viscous coefficient of mono-articular muscles around shoulder, elbow and wrist joints, respectively. The values of B_{11}, B_{22} and the off-diagonal components are determined by the measurement results (e.g., [9]). However, B_{33} which depends on the torque of wrist has never been measured. In this research, the value of B_{33} is empirically set to the mean of B_{11} and B_{22}.

$$B = \begin{bmatrix} B_{11} & B_{21} & 0 \\ B_{12} & B_{22} & 0 \\ 0 & 0 & B_{33} \end{bmatrix} = \begin{bmatrix} 0.63 + 0.095|\tau_s| & 0.175 + 0.0375|\tau_e| & 0 \\ 0.175 + 0.0375|\tau_e| & 0.76 + 0.185|\tau_e| & 0 \\ 0 & 0 & 0.695 + 0.14|\tau_w| \end{bmatrix}$$

Here, τ_s, τ_e, and τ_w stand for torque of shoulder, elbow and wrist, respectively. In order to evaluate the dependence of the model on the wrist viscosity, we also compare the result by varying B_{33} in the range of $\pm 20\%$.

3 Method

In order to analyze arm postures during reaching movement, we set up the experiment environment as shown in Fig. 1. All the subjects were 22–26 years old right-handed males. Subject's arm was floated by two air-pucks to balance frication and gravity. All the subjects were asked to move their right arms from start position to target, and all motions were measured by OPTOTRAK 3020(Northern Digital Inc.) at 200 Hz. Figure 2 shows one example of measured postures.

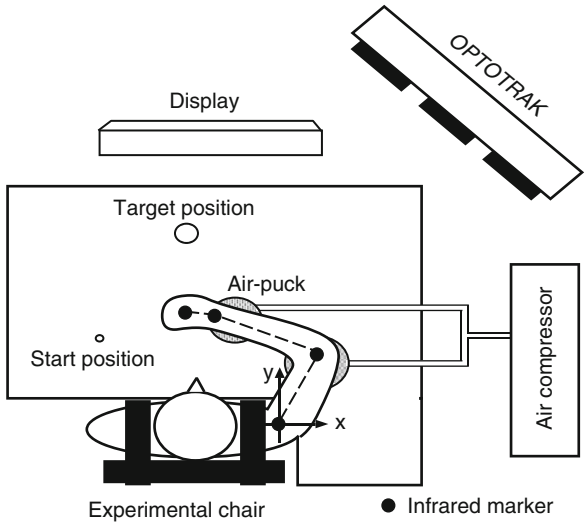


Fig. 1 Experimental setup

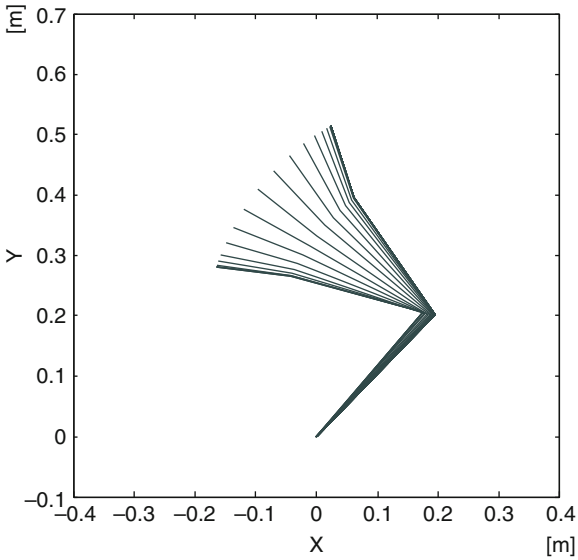


Fig. 2 Human arm postures during movement

3.1 Measurement Conditions

Subjects were divided into two groups (Groups A and B) to carry out two kinds of experiments as shown in Fig. 3. For a random nature movement condition(RN), the start and target positions were randomly selected and the distance of the movement was set to 45% of the arm length. The subjects were asked to move their arm from start position to target naturally just like them did in their daily life. For a specified quick movement condition(SQ), the start and target positions of the movements

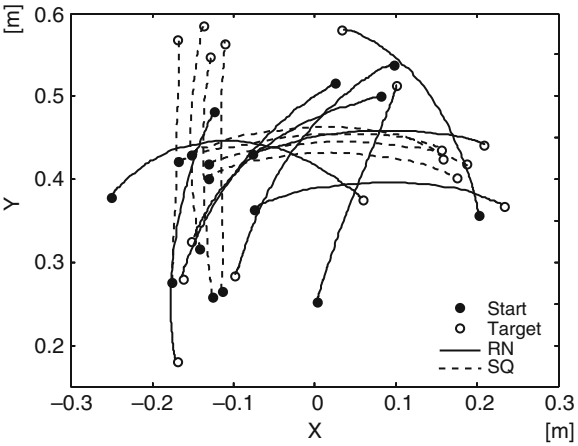


Fig. 3 Measured hand paths

were specified into two kinds. Then the movement distance was fixed to 30 cm, and the movement duration was limited within a range from 550 to 650 ms.

3.2 Filtering and Analysis

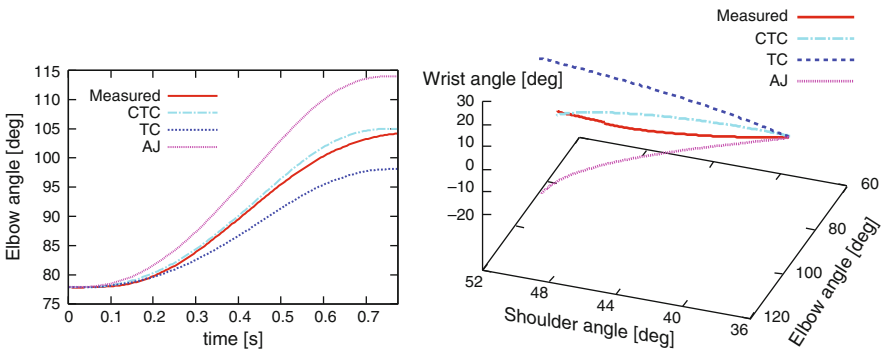
The position data of the four makers were digitally filtered by a fourth-order Butterworth filter with a cutoff frequency of 10 Hz. The actual beginning and end positions of each movement were detected using a two-dimensional curvature with a 3 mm^{-1} threshold. Outliers of measured data were also removed.

3.3 Real-Coded Genetic Algorithms (RCGA)

We built a real-coded genetic algorithm (RCGA) to calculate the optimal postures based on the above models. The movement duration was divided into six parts, such as the optimization method of Harris et al. [7], and the set of the wrist joint angles at every divided time was used as the representation of the individual of RCGA. In the calculation of the optimal solution, the hand tip trajectory was always set to be the same as the measured data, so that the optimal calculation would only focus on coordinates transformation problem.

4 Results and Discussion

Figure 4 shows a typical subject's the results of the measured postures and the optimal postures for each model. Figure 4b shows the relationship of three joints for the movement; Fig. 4a shows the elbow angular change during the movement time. Table 1 shows the average of the RMS deviations between the measured postures and the optimal ones for a whole movement duration. From these quantitatively



(a) Elbow-joint angle (b) All joints

Fig. 4 Comparison of the measured postures and the optimal postures. **a** typical subject

Table 1 Comparison of RMS deviations between measured postures and the optimal postures.[rad.] Here, α is a coefficient that $B_{33}{}' = \alpha B_{33}$, $B_{33}{}'$ stands for the B_{33} after variation

Models		Group A	Group B	All subjects
AJ		0.1388	0.1254	0.1321
TC		0.1147	0.3577	0.2362
	$\alpha = 0.8$	0.0550	0.1403	0.0976
CTC	$\alpha = 1.0$	0.0377	0.0539	0.0458
	$\alpha = 1.2$	0.0359	0.0713	0.0536

comparison, we found that the optimal arm postures that are predicted by CTC are closest to the measured arm postures. This result consists with the result of Tsuyuki et al. [8] that investigated for a 2-link arm in the three-dimensional space. Meanwhile, in order to evaluate the dependence of CTC on the viscosity, we compare the result by varying B_{33} in the range of $\pm 20\%$. As shown in Table 1, although the wrist viscosity is altered, the RMS deviations of CTC still keep lower than AJ and TC.

5 Conclusion

According to the analysis above, we note that CTC can predict the best optimal posture, and is also less dependent on the viscosity of arm. Finally, we conclude that the minimum commanded torque change model (CTC) is more plausible for the human arm posture selection during reaching movements.

References

1. Kawato, M.: Optimization and learning in neural networks for formation and control of coordinated movement. In Meyer, D.E., Kornblum, S., eds.: Attention and Performance XIV. Cambridge, MA: MIT Press (1992) 821–849.
2. Abend, W., Bizzi, E., Morasso, P.: Human arm trajectory formation. Brain **105** (1982) 331–348.
3. Flash, T., Hogan, N.: The coordination of arm movement: An experimentally confirmed mathematical criterion. J. Neurosci. **5** (1985) 1688–1703.
4. Rosenbaum, D.A., Loukopoulos, L.D., Meulenbroek, R.G., Vaughan, J., Engelbrecht, S.E.: Planning reaches by evaluating stored postures. Psychol. Rev. **102** (1995) 28–67.
5. Uno, Y., Kawato, M., Suzuki, R.: Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model. Biol. Cybern. **61** (1989) 89–101.
6. Nakano, E., Imamizu, H., Osu, R., Uno, Y., Gomi, H., Yoshioka, T., Kawato, M.: Quantitative examinations of internal representations for arm trajectory planning: Minimum commanded torque change model. J. Neurophysiol. **81** (1999) 2140–2155.
7. Harris, C.M., Wolpert, D.M.: Signal-dependent noise determines motor planning. Nature **394** (1998) 780–784.
8. Tsuyuki, K., Yamanaka, K., Soga, Y., Wada, Y., Kawato, M.: Human arm posture planning based on optimal criterion in three-dimensional space. Proceedings of the 18th Symposium of Biological and Physiological Engineering (2003) 123–124 in Japanese.
9. Gomi, H., Osu, R.: Task dependent viscoelasticity of human multijoint-arm and its spatial characteristics for interation with environments. J. Neurosci. **18** (1998) 8965–8978

Decoding the Grid Cells for Metric Navigation Using the Residue Numeral System

Cécile Masson and Benoît Girard

Abstract Knowing that the grid cells in the dorso-lateral band of the medial entorhinal cortex are a necessary component of rat's ability to perform path integration and homing, we explore the idea proposed by Fiete et al. (2008, J. Neurosci. **28**(27):6858) that they might encode the coordinates of the rat in a Residue Number System. We show that the decoding network based on a gradient descent they propose to use for that purpose does not operate satisfactorily, and thus propose a network using a direct approach derived from the Chinese Remainder Theorem.

Keywords Grid cells · Entorhinal cortex

1 Introduction

Rodents are able to directly come back to their departure point after having explored an unknown environment, even in the absence of visual or other allocentric cues, exhibiting the so-called *homing* navigation strategy [1]. They achieve this using path integration (or dead-reckoning), i.e. the integration of self-movement information derived from proprioceptive inputs, efferent copy of movement orders and vestibular information, so as to continuously estimate their position relative to their departure point.

The neural substrate of this integration mechanism is thought to be the recently discovered grid cells (GC) [2] in the dorso-lateral band of the medial entorhinal cortex (dMEC). These cells fire according to a repeating equilateral triangular pattern (Fig. 1, left) in the locomotion plane, characterized by a given period (the distance between two maximal response sites) and a given orientation (of one of the triangle sides with regard to an allocentric reference). Nearby cells in the dMEC have the same period and orientation but a different phase and are thus thought to belong to the same grid, while increasing grid sizes have been observed when recording cells

B. Girard (✉)
ISIR, UMR 7222, UPMC – CNRS, F-75252 Paris Cedex, France
e-mail: benoit.girard@isir.upmc.fr

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_73,

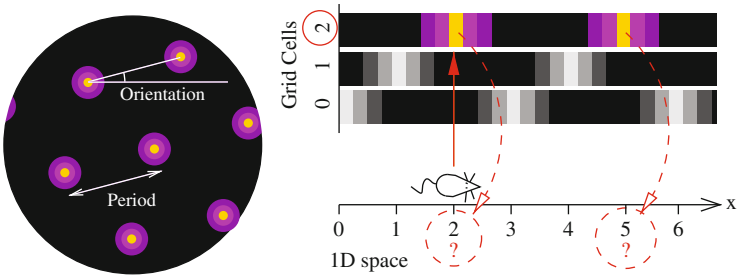


Fig. 1 *Left:* Grid cell firing pattern schema. A grid cell recorded while the rat locomotes (here in a circular arena) fires with an equilateral *triangular* pattern characterized by a period and an orientation. *Right:* Grid cells as a modulo operator. Considering the 1D case, with a grid of period 3, the current rat position is encoded (*plain arrow*) by neuron 2, but neuron 2 represents all positions *2 modulo 3* (*hashed arrows*), adapted from [8]

along a dorsal to ventral axis. This spatial pattern of activity takes into account self-movement information, as it has been shown that it is preserved in absence of any visual cue, despite a spreading of the bumps of activity, caused by accumulation of errors by the integration process.

The dMEC is an essential part of the path integration/homing system, as it has been shown that animals with an entorhinal lesion are unable to perform homing [3]. Interestingly, the hippocampus does not seem to be an essential component of path integration, as hippocampectomized rats can still perform homing [4]. Consequently, the intrinsically metric homing navigation strategy does not require functional place cells – that are a topological code of the rat position rather than a metric one – to operate. How to perform homing using the GC activity solely is a yet unanswered question, as most existing models of navigation including the GCs (like the very first one [5]) use them to generate place cells, which can then be exploited to navigate with topographical strategies, like place recognition triggered response or path planning (see [6, 7] for reviews on this matter). A computational model of homing should be able to extract directly from the grid cells the current position with regard to the departure point, so as to be able to generate the opposite locomotor command.

In the following, we propose a computational model able to extract directly from the activity of a set of grid cells the current coordinates of the animal, that can be used to control the homing behavior.

2 Model

2.1 The Initial Idea

In a recent paper, Fiete et al. [8] proposed a new way to interpret the GC activity: they could be considered as performing the 2D equivalent of modulo operator. When

considering one axis of a given grid, the currently most active neuron in this grid provides the residue of the division of the current coordinate of the rat on this axis by the grid period (Fig. 1, right). The information from one grid allows to locate the animal modulo this grid's period, but the information provided by a set of N grids of periods $(\lambda_1, ..., \lambda_N)$ corresponds to a well known number encoding system, the Residue Number System (RNS). The RNS is based on the Chinese Remainder Theorem (CRT) which states that given a set of residues $(r_1, ..., r_N)$ and a set of pairwise coprimes $(\lambda_1, ..., \lambda_N)$ (with $\Lambda = \prod_{i=1}^{N} \lambda_i$), there exist a unique integer x modulo Λ so that $\forall i \in [1, N]$, $x \equiv r_i \pmod{\lambda_i}$. Thus, given the N aforementioned λ_i any number in $[0, \Lambda[$ can be encoded as the set of residues r_i. Note that this can be generalized to non coprime pairwise periods, in which case the Λ is the least common multiple (l cm) of the periods. If we consider a set of grids of identical orientation and different periods, and use them as modulo operators, one should be able to extract the coordinate x of the animal on one of their axes. Doing this operation on two axes of these grids provides a coordinate system sufficient to locate the animal in space, as long as it does not travel distances exceeding the periods' l cm.

Fiete et al. did not provide a computational implementation of their idea, but suggested to decode the activity of a set of grids using a neural network model proposed by Sun and Yao [9]. However, this model computes x by performing a gradient descent on an discontinuous energy function with multiple local minima:

$$E(r) = \sum_{i=1}^{N} \frac{1}{2} |x \bmod \lambda_i - r_i|^2 + \frac{1}{2AR} x^2 \tag{1}$$

Such a function being inadequate for a gradient descent, it work quite poorly as it easily get stuck in local minima. For example, with a set of periods 2, 3, 5, 7, Fig. 2

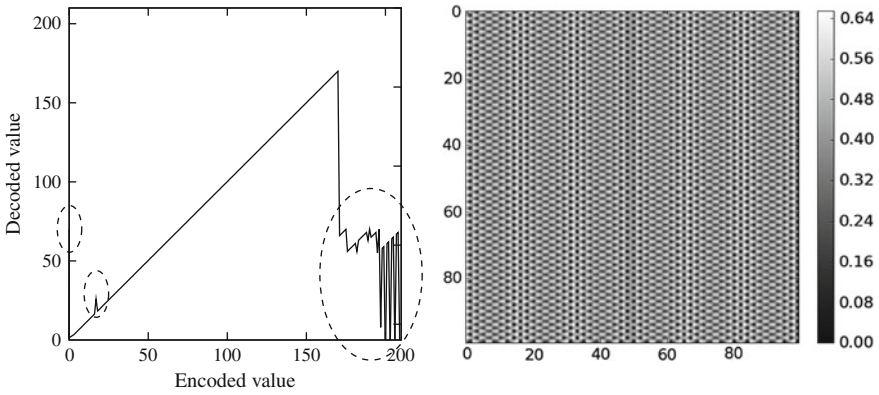


Fig. 2 *Left:* The Sun and Yao [9] model fails to decode correctly all the encoded values, errors, away from the *diagonal*, are highlighted by *hashed ellipses*. *Right:* Decoding errors (in cm) generated by our model on a 100 × 100 m surface

(left) shows the decoding of all the positions in $[0, \Lambda = 210[$, where numerous values are not correctly decoded.

2.2 TRC-Based Model

We propose a decoding scheme based on an explicit use of the basic mathematics associated to the CRT. Consider the $\hat{\lambda}_i = \frac{\Lambda}{\lambda_i} = \prod_{j \neq i} \lambda_j$: they are pairwise coprimes, and according to Bezout's theorem, u_i and v_i exist such that $u_i \lambda_i + v_i \lambda_i = 1$. If we define $e_i = v_i \hat{\lambda}_i$, x can be computed as the following weighted sum:

$$x = \sum_{i=0}^{N} e_i r_i \tag{2}$$

A similar solution holds when using non-pairwise coprime periods.

We built a computational model to show the efficiency of the Fiete et al. RNS idea, using the direct RNS computation (Eq. 2). This model (Fig. 3) operates in 2D and uses real residue values extracted from grid cells. It uses velocity inputs provided to 4 grids , based on the model proposed in [10]. These grids have realistic periods of 38, 50, 62 and 74 cm, theoretically allowing the unique decoding of values over more than 5 km; they all have the same orientation, which also seems to be realistic [11]. The position of the maximum activity on two of the three axes of the grids is obtained by summing the activity of all cells projected on these axes and computing the circular barycenters of the resulting vectors. These values correspond the residues r_i, which are then sent to two neurons, whose synaptic weights were set to e_i.

The simulations performed with this model show that a simulated rat is able to compute its position in 2D. The errors, caused by the discretization of the grids, have an average of 0.39 cm ($std = 0.19$) (Fig. 2, right), a value which seems acceptable with regards to the size of a rodent.

3 Discussion and Conclusion

Based on the idea – proposed by Fiete et al. [8] – that the grid cells could be considered as an encoder of the position of animal using a residue number system, we propose a computational model performing the decoding of this position, to be used as a command for the locomotor system in a homing navigation strategy. We also establish that the residue-to-decimal decoder proposed in [9] does not operate efficiently.

Our simulations were conducted without noise in the speed measurements. In reality, such noise affects the grid updating and is fought with resets of the grid activity in places identified by learnt allocentric cues [2]. Such stabilization mechanisms are present in most GC models, and are either based on direct visual inputs

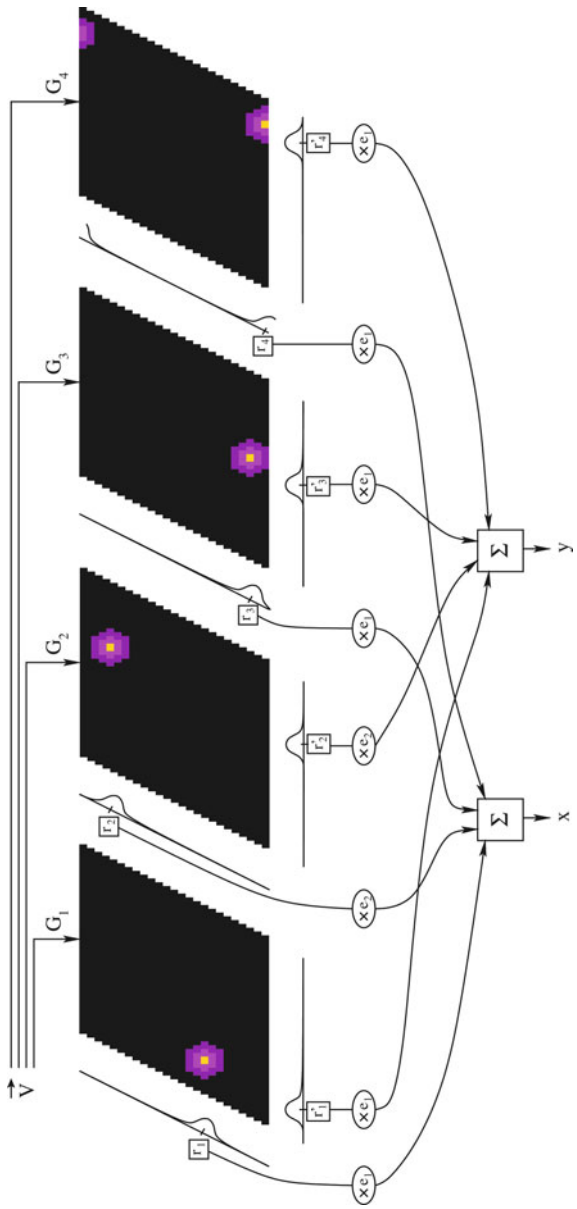


Fig. 3 Grid cells decoding model. The speed V of the simulated animal displaces the activity in the four grids (25×25 neurons), the gains G_i define the grid periods (38, 50, 62 and 74 cm). The circular barycenters r_i and r'_i of the grid activities projected on one of their axes are the residues used with the e_i weights to decode the coordinates (x, x') (note that this reference frame is not orthogonal)

or on interactions with the multi-modal place cells [12]. This does not affect the validity of proposed decoding scheme, which should of course be connected to a grid cell model including a stabilization mechanism.

Finally, a computational model – anterior to the discovery of the grid cells [13] – proposed to learn the coordinates corresponding to each place cell of a hippocampus model, so as to thereafter enable metric navigation. We showed here that the grid cells, providing inputs to the place cells, are sufficient to perform metric navigation, with the advantage that no prior associative learning phase is necessary, in accordance with observations of homing behavior performed by rodents in new and unexplored environments.

Acknowledgements Cécile Masson acknowledges the support of the EC Framework 6 ICEA project (Integrating Cognition, Emotion and Autonomy), IST-4-027819-IP

References

1. Etienne, A., Jeffery, K.: Path integration in mammals. Hippocampus **14**(2) (2004) 180–192
2. Hafting, T., Fyhn, M., Molden, S., Moser, M., Moser, E.: Microstructure of a spatial map in the entorhinal cortex. Nature **436** (2005) 801–806.
3. Parron, C., Save, E.: Evidence for entorhinal and parietal cortices involvement in path integration in the rat. Exp. Brain Res. **159** (2004) 349–359.
4. Alyan, S., McNaughton, B.: Hippocampectomized rats are capable of homing by path integration. Behav. Neurosci. **113**(1) (1999) 19–31.
5. Fuhs, M., Touretzky, D.: A spin glass model of path integration in rat medial entorhinal cortex. J. Neurosci. **26**(16) (2006) 4266–4276.
6. McNaughton, B., Battaglia, F., Jensen, O., Moser, E., Moser, M.: Path integration and the neural basis of the'cognitive map'. Nat. Rev. Neurosci. **7**(8) (2006) 663–678.
7. Moser, E., Kropff, E., Moser, M.: Place cells, grid cells, and the brain's spatial representation system. Neuroscience **31** (2008) 69–89.
8. Fiete, I., Burak, Y., Brookings, T.: What grid cells convey about rat location. J. Neurosci. **28**(27) (2008) 6858
9. Sun, H., Yao, T.: A neural-like network approach to residue-to-decimal conversion. IEEE Int. Conf. Neural Netw. **6** (1994) 3883–3887.
10. Sheynikhovich, D.: Spatial navigation in geometric mazes : A computational model of rodent behavior. PhD Thesis, EPFL (2007)
11. Barry, C., Hayman, R., Burgess, N., Jeffery, K.: Experience-dependent rescaling of entorhinal grids. Nat. Neurosci. **10**(6) (2007) 682–684.
12. Samu, D., Erős, P., Ujfalussy, B., Kiss, T.: Robust path integration in the entorhinal grid cell system with hippocampal feed-back. Biol. Cybern. (2009)
13. Foster, D., Morris, R., Dayan, P.: A model of hippocampally dependent navigation, using the temporal difference learning rule. Hippocampus **10**(1) (2000) 1–16.

Discussion on Rhythmic Gait Movement Affected by Cerebral Cortex Signal

Wei Dong, Rubin Wang, Zhikang Zhang, and Jianting Cao

Abstract The gait movement on humans is a typical mode of the rhythmic movement. Proved by experiment, there is a neural circuit existing in the body of creatures, which controls various rhythmic movements. It is called as central pattern generator. The researchers study the gait movement in the basis of establishing the CPG model now. For the rhythmic movement, the neurobiologists believe that it is a lower neural activity, which is a spontaneous behavior without the control of the cerebral cortex signal. However, the gait movement on human has some different modes. Every mode of the gait movement is not directly controlled by cerebral cortex signal, but the instruction signal generated by cerebral cortex can regulate the interchange between the gait modes. In this paper, in the basis of thinking about the instruction regulation of cerebral cortex to gait movement on humans, the CPG model is revised. The revised model can reflect that the cerebral cortex signal can regulate the mode and frequency of gait, and can realize the interchange between gaits.

Keywords CPG · Gait movement · Rhythmic movement · Cerebral cortex signal · Conversion function

1 Introduction

The essential purpose of neuroscience is to explain how the nervous system brings movement. Neurobiologists generally believe that the rhythmic movement is nothing to do with brain awareness, but self-excided action of lower neural center, a space-time motion mode produced and controlled by central pattern generator (CPG).

W. Dong (✉)
School of Information Science and Engineering, Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, P.R. China
e-mail: 020060098@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_74,

In order to study the various functions of CPG, the researchers have established some related mathematical models. These mathematics models are useful and convenient for the engineering application. Some researchers also further studied some important properties of CPG, such as the effect of tonic input and sensory feedback, robustness, stable oscillation and coupling relationship. Through the simulation study, we can well understand the CPG with a comprehensive point of view, and provide firm foundation for CPG application in engineering and medicine rehabilitation. In this paper, in basis of the fact in Biology, the CPG model was modified so as to make the simulation results of model closer to the fact [1, 2].

2 Biology CPG Model

In human, the walking movement is a typical rhythmic one. In engineering, CPG neural circuits can be taken as a distribution system composed of a group of non-linear oscillators mutually coupled, generating rhythmic signals through phase coupling, and the change of coupling relationship between oscillators can generate space-time sequence signals with different phase relationship so as to achieve different movement modes. In this paper, it is shown to set up a network model of the CPG based on the neural oscillator theory, depending upon the simplified leg muscles model, so as to simulate the rhythmic movement mode generated when gait movement [3].

In Fig. 1, the two groups of CPG meaning two legs are connected each other. In one leg, there are three couples of neural oscillator units, meaning three joint-hip,

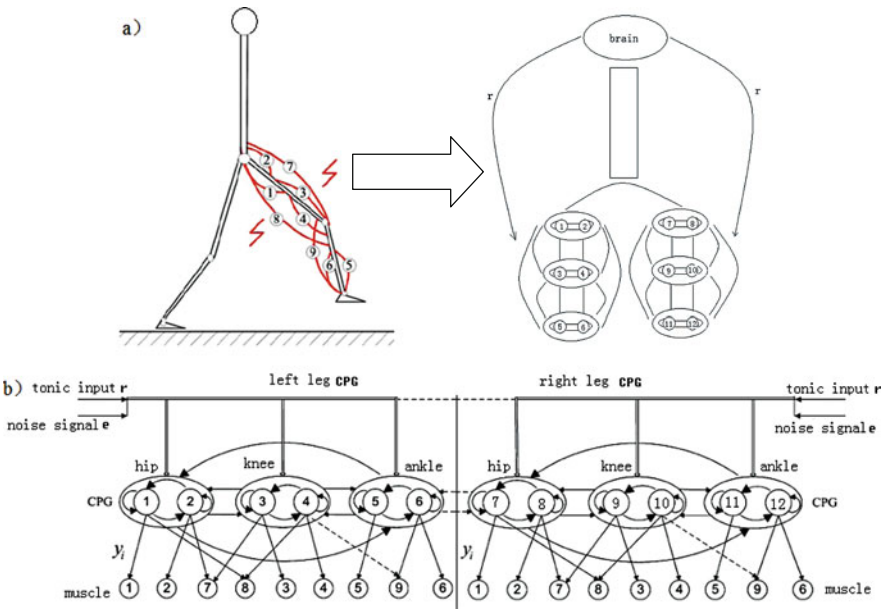


Fig. 1 The leg CPG model (**a**. the relationship between CPG model and leg structure; **b**. the CPG network)

knee and ankle. There are two neuron groups in one couple of oscillator unit, meaning two kinds of muscle groups-flexor and extensor [4].

3 Modified CPG Model

Our research focuses on the variety property of rhythmic gait movement. For the theoretic analysis on "variety" property, it has been introduced in the previous research [2]. Nowadays, simulation study on the output modes of legs CPG model only limited to study movement state with the single mode, it cannot realize the interchange between the different modes, meaning it cannot show the instruction regulation function of cerebral cortex to rhythmic movement. The other papers have introduced different modes of the output of rhythmic movement. There are two examples in Fig. 2.

In Fig. 2, it indicates that, in the simulation of CPG model the different gait movement states can correspondingly generate different rhythmic output modes. We modified and improved the previous CPG model to realize gait conversion, the equation is as follows:

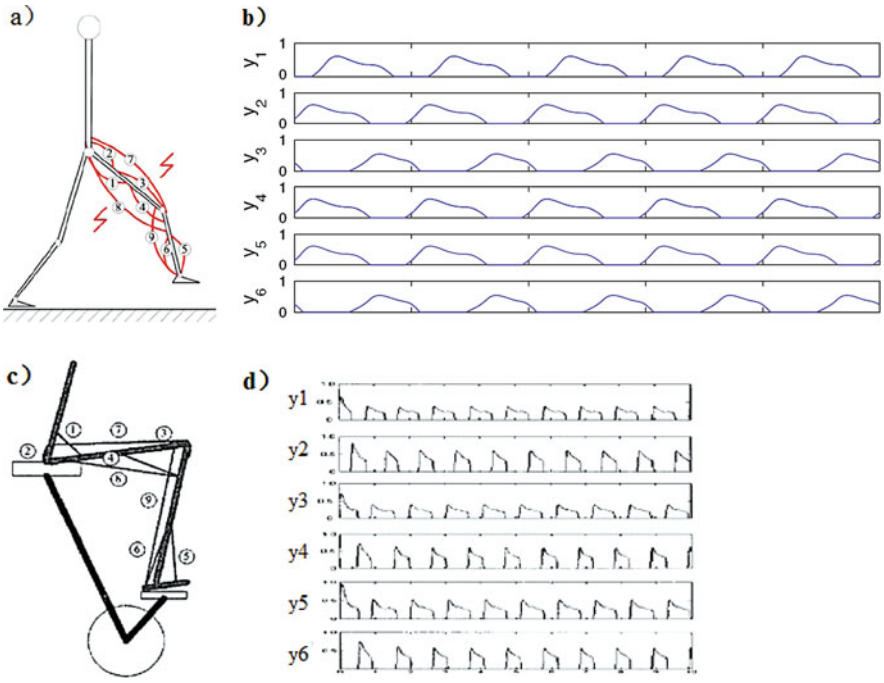


Fig. 2 (**a**) muscle skeleton structure model when walking; (**b**) the output of CPG model in simulation when walking; (**c**) muscle skeleton structure model when cycling; (**d**) the output of CPG model in simulation when cycling

$$
\begin{aligned}
&\tau \dot{X}_i = -X_i - \beta v_i + r + \sum_{i \neq j} \omega_{ij} Y_i + Le \qquad i,j = 1,2,\cdots,12 \\
&{}^{*}r = \begin{cases} r_0 \\ r_0 + r_{brain} \end{cases} \quad (t \in [t_{start}, t_{end}]) \\
&{}^{*}r_{brain} = A\ sin\left(\frac{2\pi}{T_r} t\right) \\
&{}^{*}\tau = bT_r \int r_{brain} dt + \tau_0 \\
&T \dot{v}_i = -v_i + Y_i \\
&{}^{*}T = \left({}^{c}/_{T_r}\right) \int r_{brain} dt + T_0 \\
&Y_i = f(X_i) = \max(X_i, 0)
\end{aligned}
\tag{1}
$$

Here, i, j mean the 1th–12th neuron; X_i means the two states of neuronal groups in the ith joint – flexor and extensor; v_i, the adaptability parameter; ω_{ij}, interconnection weight between the 12 neuronal groups of both legs; τ is status constant; T is fitness constant. As for the parameters in detail, please refer to literature [5].

For our modified model, it is emphasized to the regulation function and the effect of tonic input r to other parameters in the CPG network, so as to realize that the CPG output modes can transform between two modes only depending on regulating the tonic input r without changing the other parameters, making it closer to the stimulus regulation in the fact.

4 Simulation Results

In the simulation, we take the constant signal as the stimulus when self-oscillation, and take the sine wave as the instruction stimulus of cerebral cortex [6]. When the cerebral cortex send out the instantaneous instruction signal, the CPG output modes can occur to change, so as to realize the regulation function of cerebral cortex.

(1) *Interchange between different modes*: Here, we only give two examples, which respect two kinds of different CPG output modes. It is shown to the interchange between these two modes in Fig. 3. There are three interchange relationships on CPG output mode.

Condition I: When the sine wave meaning the instruction signal from cerebral cortex only is the state of positive semi-period, CPG output mode can transform from type I to type II (Fig. 3).

Condition II: When the sine wave meaning the instruction signal from cerebral cortex only is the state of negative semi-period, CPG output mode can transform from type II to type I (Fig. 3).

Condition III: When the sine wave meaning the instruction signal from cerebral cortex is the state of a whole period, CPG output mode cannot transform.

(2) Interchange between different frequencies in the same mode: As for the regulation signal from cerebral cortex r_{brain}, it not only can make the different CPG modes interchange, but also can make the different frequencies in the same CPG

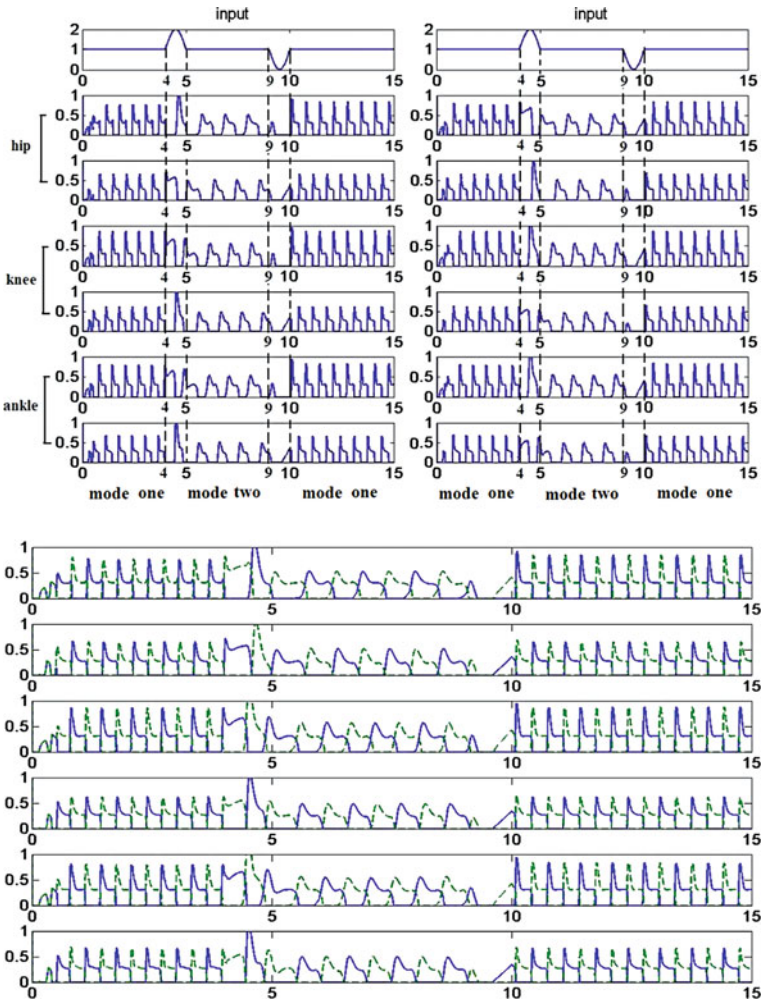


Fig. 3 Interchange between modes

mode interchange. There is a close relationship between the cycle T_r of the regulation signal from cerebral cortex r_{brain} and the time parameter τ and T in the CPG network, adjusting to the effect parameter b, c, the interchange between the different frequencies in the same mode can be realized in the simulation. Accordingly, there are three conditions:

Condition I: When the sine wave meaning the instruction signal from cerebral cortex only is the state of positive semi-period, CPG output frequency in the same mode can transform from fast to slow (Fig. 4).

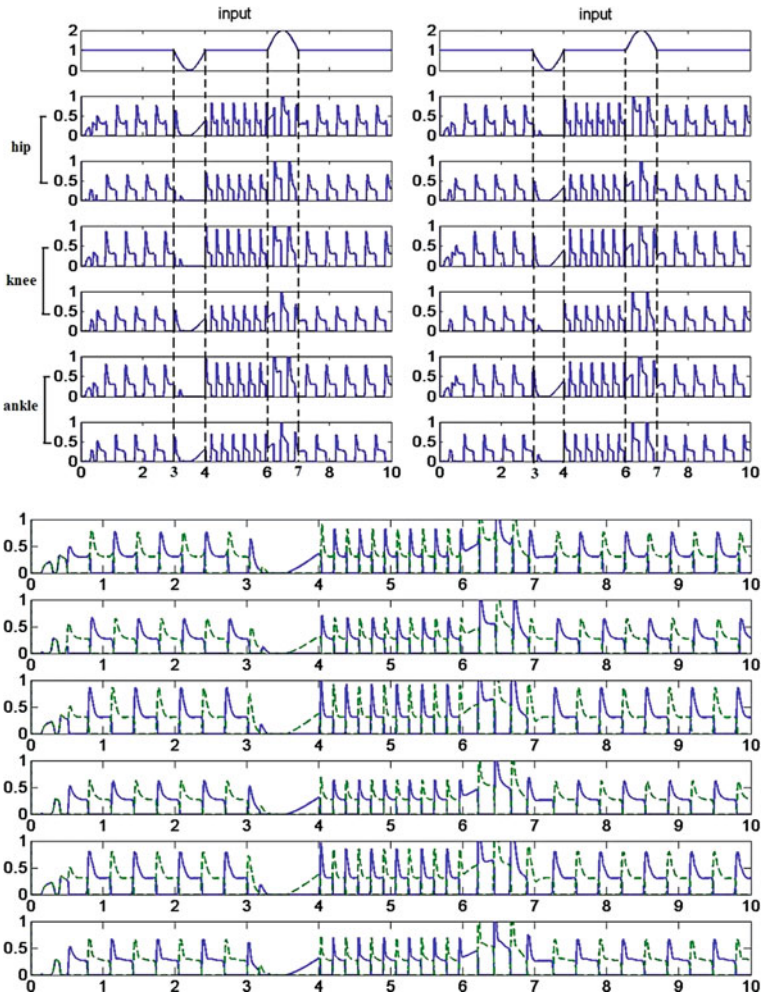


Fig. 4 Interchange between frequencies in the same mode

Condition II: When the sine wave meaning the instruction signal from cerebral cortex only is the state of negative semi-period, CPG output frequency in the same mode can transform from slow to fast (Fig. 4).

Condition III: When the sine wave meaning the instruction signal from cerebral cortex is the state of a whole period, CPG output frequency in the same mode cannot transform.

The modified model can better show the behavior transformation function, it not only can interchange different frequencies in the same mode but also can interchange different modes. This modifiability better improves the theoretic meanings of CPG model itself, and makes simulation study theoretically closer to the fact.

Acknowledgements This work is supported by National Nature Science Foundation of China (10672057, 10872068).

References

1. Matsuoka, K.: Mechanisms of frequency and pattern control in the neural rhythm generators. Biol. Cybern. **56** (1987) 345–353.
2. Wei, D., Rubin, W., Zhikang, Z.: Simulation study of CPG model: Exploring a certain characteristics of rhythm of gait movement on the intelligent creature. Sixth International Symposium on Neural Networks (2009).
3. Zhang, D., Zhu, K., Zheng, H.: Model the leg cycling movement with neural oscillator. IEEE International Conference on Systems, Man and Cybernetics (2004) 740–744.
4. Zhang, D., Zhu, K.: Computer simulation study on central pattern generator: from biology to engineering. Int. J. Neural Syst. **16**(6) (2006) 405–422.
5. Wei, D., Rubin, W., Zhikang, Z.: The simulation study on the pattern of muscles controlled by CPG in rhythm gait movement. J. Dyn. Control. **6**(4) (2008) 327–331(Chinese Journal).
6. Todorov, E.: Cosine tuning minimizes motor errors. Neural Comput. **14** (2002) 1233–1260.

Re-optimization Contributes to the Adaption of External VF Field

Chunjiang Fu, Rubin Wang, and Jianting Cao

Abstract The computational model of Franklin et al. (J. Neurosci., **28**(44) 2008, 11165–11173) describes the adaptation of human arm movement to the external force field with lower-level feed-forward module for compensation. While this paper proposes another possible option that when the environment can be expressed in the internal model, the upper-level optimal control module can directly play the role of feedforward compensation merely via tuning two corresponding parameters: the co-contraction level r between a pair of antagonistic muscles and the parameter reflecting the force field during optimal calculation. To verify this idea, this paper imposes velocity-dependent force field in the simulation on the neruo-musculo-skeletal biomechanical model using optimal control method. Qualitatively, the outcome is consistent with the experimental references.

Keywords Motor adaption · Velocity-dependent force field · Feedfoward compensation · Neuro-musculo-skeletal biomechanical model · Optimal control

1 Introduction

Voluntary movement, known for its complexity, flexibility, stability and adaptability, has been extensively researched and will still be on the cutting edge in future. The exploration for its bio-control mechanism would undoubtedly benefits related subjects like robotic technology, medical rehabilitation and sports training. Recently, numerous studies focus on the motor learning and adaptation with reaching paradigm, which needs movement plan [1–3] and stability maintenance [4–6].

C. Fu (✉)
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, China
e-mail: fcj2@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_75,

Franklin et al. [6] proposes a concise and effective model. However, its planned movement is derived from experimental data fitting which implicitly excludes that the feedforward learning and adaptation process may happen in the upper-level during movement planning calculation through optimal control algorithm. In contrast, this paper revises Franklin et al. [6] model and attributes the external force field compensation to the upper-level optimal control module. In addition, some other revision contains nonlinear biomechanical arm model instead of linear model in Franklin et al. [6] and adjusting only two upper-level parameters rather than many fitting parameters without obvious biological meaning in Franklin et al. [6].

2 Materials and Model

In order to compare with the original paper, this paper follows the experimental setup in Franklin et al. [6]. The hand of subjects who receives tests is fixed with the PFM robotic device on the horizontal plane, doing outward reaching movement toward the target ahead within a span of time about 0.6 s. At the same time, the PFM exerts a velocity-dependent force field according to the hand movement speed v and the exact expression is $Fe = -[13–18;\ 18\ 13] \times v$. Furthermore, the scale of the amplitude is set to 2/3Fe because of different individual. Motion range is from (0, 0.3) m to (0,0.55) m taking the right shoulder as the origin.

With the description in introduction, the whole simulation model of this paper is presented in Fig. 1. The internal model with the form of optimal control carries out internal simulation before the real execution by arm. During this process, the internal model revise the muscle co-activation level r for human limb and identify the external environment through the parameter f, simply in this paper, the presentation of external force field is multiplying a coefficient f to Fe and apply the product

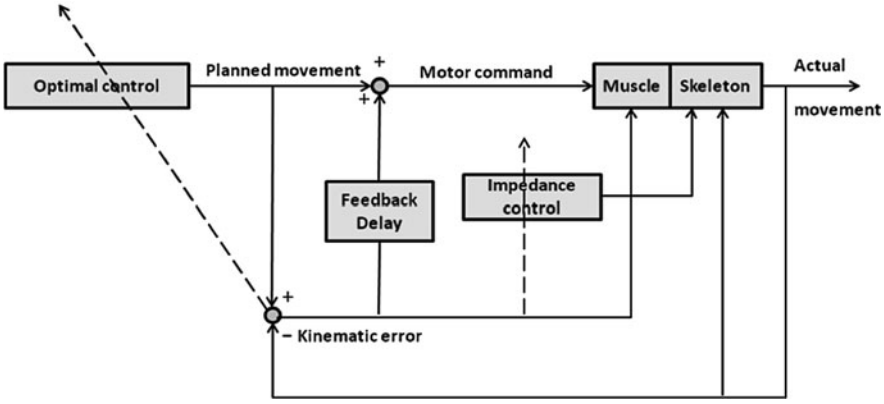


Fig. 1 Control diagram of this paper

in the optimal calculation using Tomlab software. However, the other parts of the simulation are operated by Matlab/Simulink.

Except the most feature of moving the adaptation module to the upper-level of optimal control block, the muscle-skeletal biomechanical model is also changed to nonlinear one from Kashima et al. [7].

The minimum criterion for optimal control is minimal motor neural signal.

Finally, Physiological parameters is chosen in line with Franklin experiment and the impedance module detail can be found in Fu et al. [8].

3 Results

Figure 2 is the comparison chart. On the left is Franklin et al. [6] experimental results and on the right are the simulation outcomes of this paper. Dotted line represents the hand trajectory, three phases consists arm movement under no external force field (NF, see A), adaptation when exposed to external velocity-dependent force field (VF, see B) and consequent straight trajectory after trained (see C).

The most obvious feature of the adaptive process is that the extent of bending to left is gradually decreasing. In this paper, we adopt two parameters involved in the optimal calculation in the upper-level internal model. The tuning process is first increase co-contraction level r from 0.1 to 0.4, which means increase the co-activation level thus augment the overall impedance of all pair of muscles, the effect is slightly reduce the trajectory curve; Secondly, hold r in constant and raise f parameter from 0 to 0.4, the intuitive impact is the plan movement trajectory first bend to left and then to right, which could compensate the influence of VF exerting on hand shaping the trajectory first bending to right and then to left. Thus the larger f is, the more compensation could be made and appropriate compensation would bring about straighter actual trajectory ; the third phase is somewhat interesting that larger f only does not provide finer outcome, at the same time, we should lower the value of co-activation r, or we may obtain over-compensated consequences, the

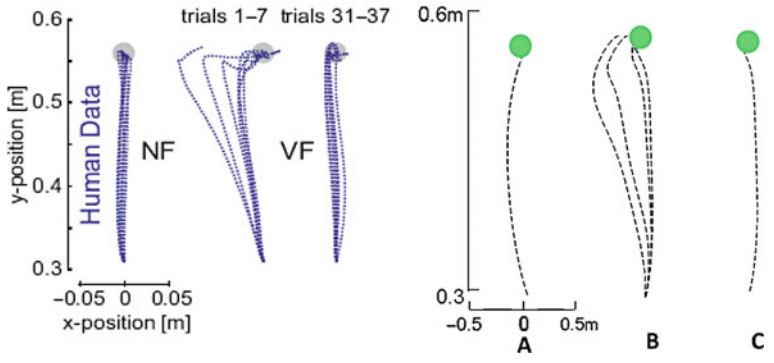


Fig. 2 Experimental results vs. simulation outcomes

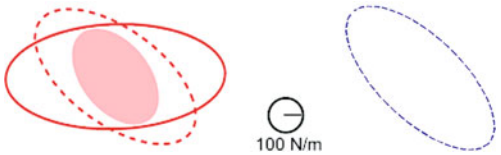


Fig. 3 Stiffness ellipse after adaption

trajectory with the adaptive parameters $r = 0.4$, $f = 0.4$ is over-compensated and a little bit jittering, not so good as another group of parameters $f = 0.5$, $r = 0.2$. The adaptive process described above is coincident with the experience of Franklin et al. [5].

The above argument is mainly focus on the movement plan aspect, on the other hand, impedance control and adaption also plays an irreplaceable role in overcoming external interference, time delay and internal nerve multiplicative noise. In this paper, the value of feedback delay is set to 60 ms; noise parameters are the same as Harris and Wolpert [9]. Corresponding ellipse of adapted impedance (stiffness mainly) is shown on the right of Fig. 3, comparable to the experimental drawing on the left (only compare to the dotted ellipse under velocity-dependant force field).

4 Conclusion

This paper recognizes that feedforward compensation may be realized by adjusting only few upper-level parameters. Furthermore, the traditional belief of the existence of reference trajectory is exactly the result of open-loop optimal control under certain situation that does not require feedback. After exerted external force field, any curve trajectory is also optimal just under the current upper-level parameter while straight trajectory should not be a must.

Applying optimal control to solve human movement problem is quite a mature method, however, the mechanism of motor learning and adaption under this frame remains unclear. Various theories have been put forward, waiting for further experiments to test and update them.

Acknowledgements Project (10872068, 10672057) supported by National Natural Science Foundation of China (NSFC).

References

1. Shadmehr, R., Mussa-Ivaldi, F.A.: Adaptive representation of dynamics during learning of a motor task. J. Neurosci, **14** (1994) 3208–3224.
2. Todorov, E.: Optimality principles in sensorimotor control. Nat. Neurosci. **7** (2004) 907–915.
3. Izawa, J., Rane, T., Donchin, O., Shadmehr, R.: Motor adaptation as a process of reoptimization. J. Neurosci. **28** (2008) 2883–2891.
4. Burdet, E., Tee, K.P., Mareels, I., Milner, T.E., Chew, C.M., Franklin, D.W., Osu. R., Kawato, M.: Stability and motor adaptation in human arm movements. Biol. Cybern. **94** (2006) 20–32.

5. Franklin, D.W., Burdet, E., Osu, R., Kawato, M., Milner, T.E.: Functional significance of stiffness in adaptation of multijoint arm movements to stable and unstable dynamics. Exp. Brain Res. **151** (2003) 145–157.
6. Franklin, D.W., Burdet, E., Tee, K.P., Osu, R., et al.: CNS learns stable, accurate, and efficient movements using a simple algorithm. J. Neurosci. **28**(44) (2008) 11165–11173.
7. Kashima, T., Isurugi, Y., Shima, M.: Control characteristics of a neuromuscular system in reaching movements. SICE (2004).
8. Fu, C., Wang, R.: The influence of co-contraction to the arm impedance during free planar movement. Procdcedings of the 3rd International Conference on Bioinformatics and Biomedical Engineering. ICBBE (2009).
9. Harris, C.M., Wolpert, D.M.: Signal-dependent noise determines motor planning. Nature **394** (1998) 780–784.

Vector Sum of Two Eyes' Images Might Be an Important Cue in Human Visual System

Qin Wang and Lei Wang

Abstract In this paper, we inferred that vector sum of two eyes' images acts as an important cue in the mechanism of human vision. In our recent report on the perceptual properties of an approaching object in periphery visual field, it was found that the perceptual performance was remarkably worse for the object moving toward the middle point of two eyes even moving distance was several times longer and it takes V curve for the orientation change. Surprisingly, when the approaching object was occluded for one eye or fell on the blind spot of one eye, visibility of an approaching object was increased even though the case when the object moving toward the middle point of two eyes; it coincided with Tyler's result but that was found in central visual field. Considering these results and the several observations in connection with Tyler's results, we inferred that the vector sum of two eyes' image motion might be an important factor for detecting the approaching object even in fully visual field; furthermore, we inferred that vector sum of two eyes' images might be one of important cues in the mechanism of human vision.

Keywords Vector sum of two eyes' images · Approaching object · Visual field · Periphery visual field · Fully visual field

1 Introduction

Perceiving an approaching object, as an indispensable function in human vision, has been studied in concerning many fields such as vision science, traffic, and sport science and so on. Many studies have reported that changing disparity (CD) and interocular velocity difference (IOVD) cues play roles in motion-in-depth perception in central visual field; changing disparity is a change in binocular disparity over

Q. Wang (✉)
Graduate School of Information Systems, The University of Electro-Communications, Tokyo 182-8585, Japan
e-mail: qinwang@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_76,

time, interocular velocity difference refers to a difference in velocity between the two eyes' images [1–3]. In this paper, we inferred that the vector sum of two eyes' image motion might be an important factor for detecting the approaching object even in fully visual field, furthermore, we inferred that vector sum of two eyes' image might be one of important cues in human visual system.

In our recent report on the perception of an approaching motion in periphery visual field [4], it was reported that the perceptual performance was remarkably worse for the object moving toward the middle point of two eyes even moving distance was several times longer and it takes V curve for the orientation change. Surprisingly, when the approaching object was occluded for one eye or fell on the blind spot of one eye, visibility of the approaching object was increased even though the case when the object moving toward the middle point of two eyes. This phenomenon can not be given convincing interpretation according to conventional studies. In our recent reported, it was suggested that vector sum of two eyes' motion might be used to detect the approaching motion in periphery visual field. The phenomenon found in periphery visual field coincided with Tyler's result but that was found in central visual field.

What are the perceptual properties of the approaching object in central visual field? We observed the approaching object in central visual field for different moving orientations. The results were coincident with those findings in periphery visual field. According to these evidences, we inferred that the vector sum of two eyes' image motion might be an important factor for detecting the approaching object even in fully visual field. We believed that these evidences could provide one clue for clarifying the mechanism of curious phenomena such as cross-traffic accidents at intersection without any obstacles and difficulty in playing catch for a ball approaching toward the middle point of two eyes. Furthermore, we believed that vector sum of two eyes' image might be one of important cues in human visual system.

2 Perceptual Performance of an Approaching Object

2.1 Perceptual Performance of an Approaching Object in Periphery Visual Field

In the recent report on the perceptual properties of the approaching object in periphery visual field, it was viewed stereoscopically to subjects with a 3D model consisting of a number of stationary bright dots, a fixation point and optionally a single moving object. It was investigated the perceptual properties of the approaching object with the object visible for both eyes. Figure 1a shows the results of perceptual performance for different moving orientation. Percentage of correct (correct-rate) was plotted as a function of object motion direction factor (Df = C/D). The perceptual performance took V curve for different moving orientation: the performance was decreased when Df changed from Df < 1 to Df = 1, while it was increased when Df changed from Df = 1 to Df > 1; and the performance became remarkably worse

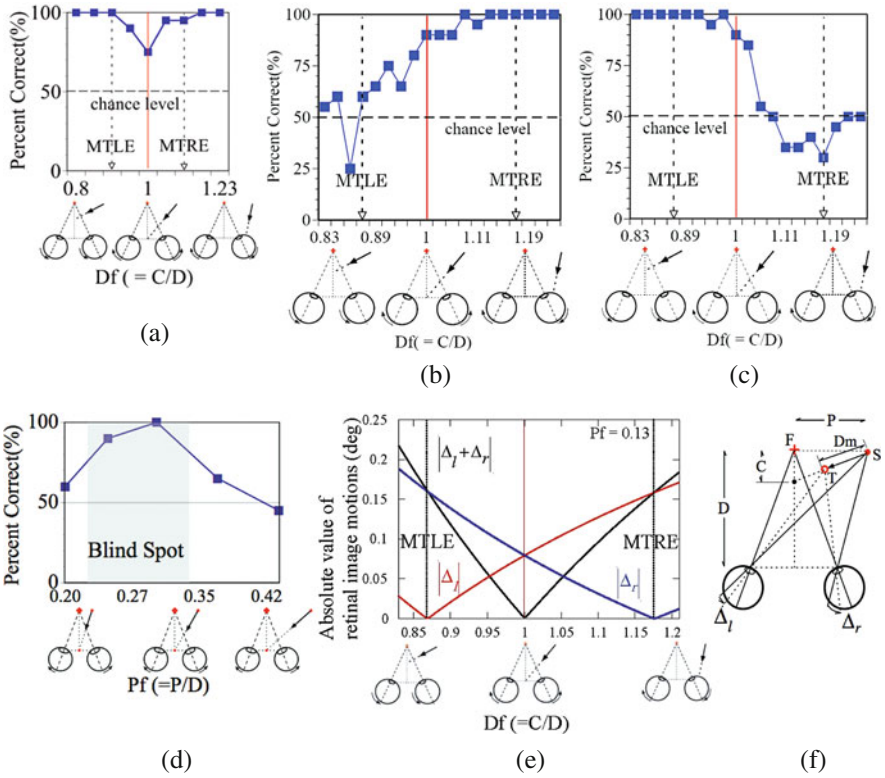


Fig. 1 Experiments of visibility of object in approaching motion for periphery visual field. (S: starting point of target stimulus. T: terminal point of target stimulus. F: fixation point. SD: stationary dots. P: distance between the fixation point and the target stimulus. D: observation distance. C: the distance between the fixation point and the intersection point which was between the extension trajectory of target motion and the mid-perpendicular of two eyes. Dm: moving distance. Df: motion direction factor of object (Df = C/D). Pf: object position factor (Pf = P/D). Δl: the resultant angular motion of the object image on the left retina. Δr: the resultant angular motion of the object image on the right retina. Df = 1: direction toward the middle point of two eyes. MTLE: direction toward left eye. MTRE: direction toward right eye.) (**a**) is the results with binocular viewing. (**b**) is the results with the object's information occluded for the right eye. (**c**) is the results with the object's information occluded for the left eye. (**d**) is for the case when the objects were inside the blind spot of on eye. (**e**) is absolute value of the summation of retinal image motions with sign (|Δl + Δr|) and the absolute value of each retinal image motions (|Δl | and |Δr|). (**f**) is the parameters of the stimuli

when Df = 1. Correct-rate depended on the orientation of object, and it became lowest in the direction toward the middle point of two eyes. That is, the motion of the object was not perceived well when the object was moving toward the middle point of two eyes.

It was investigated the perceptual properties of the approaching object with the object absent of one eye's image. Figure 1b, c show the results for the experiment in

which the stationary stimuli were visible for both eyes, but the approaching object was occluded for one eye. Figure 1d shows the results for the experiment in which the object image was falling on the blind spot of one eye. It was reported that the motion of the approaching object could be clearly perceived even in the orientation of object toward the middle point of two eyes (Df = 1) in the absent condition of one eye's image, while it could not be perceived with both eyes' images.

Figure 1e shows the absolute value of the vector summation of the image motions on the left and right retinas ($|\Delta 1 + \Delta r|$) and the absolute values of image motion on both eyes ($|\Delta l|$ and $|\Delta r|$). The vector sum is nearly zero when the object approaches toward the middle point of two eyes, which indicates having a strong correlation with the perceptual performance (Fig. 1a). In addition, $|\Delta l|$ and $|\Delta r|$ imply the correlation with the perceptual characteristics when one eye's information was absent (Figs. 1b, c). That is, the visibility of the approaching object dropped down worst for the approaching object towards the mid point of the two eyes, and it strengthened when the motion direction of the approaching object was towards one eye or not towards the midpoint of the two eyes even towards one eye.

2.2 Perceptual Performance of an Approaching Object in Central Visual Field

Tyler reported that two eyes are less sensitive than one to motion perception in the condition of the approaching object motion to the midpoint of the two eyes under the central visual field: it was corresponding to the object moving to right in front of two eyes. This phenomenon was identical with that for the special condition (Df = 1) of the approaching motion to the midpoint of the two eyes in periphery visual field in the recent report.

We observed the approaching object directing in different orientations for different visual field. The results were identical with the recently report on the perceptual properties of the approaching object in periphery visual field. That is, the visibility of the approaching object changed as V curve even in whole visual field; the visibility was worst when the approaching object moved towards the midpoint of the two eyes, and it became strengthened as direction depart from towards one eye or not towards the midpoint of the two eyes. Based these evidences, we inferred that vector sum of two eyes' image motion might be an important factor for detecting the approaching object in the fully visual field.

3 Conclusion

In this paper, we introduced the recent report on the perception of the approaching object in periphery visual field. Comparing with our observation in different visual field and Tyler's results for the special condition, we inferred that the vector sum of two eyes' image might be an important factor for detecting the approaching

object in the whole visual field. We believed that these evidences could be used for clarifying the mechanism of curious phenomena and reducing the cross traffic accident. In addition, it is expected to apply in sport science. Furthermore, we inferred that vector sum of two eyes' images acts as one of important cues in human visual system.

References

1. Brooks, K.R., Leland, S.: Stone. Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. J. Vision **4** (2004) 1061–1079.
2. Shioiri, S., Saiho, H., Yaguchi, H.: Motion in depth based on inter-ocular velocity differences. Vis. Res. **40** (2000) 2565–2572.
3. Cumming, B.G., Parker, A.J.: Binocular mechanisms for detecting motion-in-depth. Vis. Res. **34** (1994) 483–495.
4. Wang, L., Idesawa, M., Wang, Q.: A study on approaching motion perception in periphery with binocular viewing, Proceeding of International Conference of Instrumentation, Control and Information Technology (SICE Annual Conference 2008) (2008) 196–199.

A Spatial Temporal Integration Model of Primary Visual Cortex Concerning Directional Selectivity

Jianpeng Zhang, Rubin Wang, Zhikang Zhang, and Jianting Cao

Abstract Vision is the most crucial function with which human beings interact with the environment, and visual information accounts for more than 80% of the total information that human receive from the outside world. The investigation of the mechanism of visual processing is closely correlate with brain′s other functions like learning, associate memory, abstract thinking and so on. The research on the receptive field (RF) is an important component of exploring the brain mechanism of processing visual information. Recently, more and more interests have being put into the modulating effect exerted beyond classical RF. It is assumed that the modulating effect based on horizontal connectivities and higher cortex feedback might involve with contextual processing of visual representation. In this article, we proposed a simplified spatial temporal integrating model of primary visual cortex concerning directional selectivity, via simulation on the sequences of continuously changing gray scale images, we assumed the plausible contribution of the spatial temporal integration in suppressing noises and correctly representing the motion information of moving scenes.

Keywords Visual model · Motion perception · Receptive field · Primary visual cortex

1 Introduction

Visual motion perception is not only one of the most important functions for human beings, but also a breakthrough point for the investigation on the mechanisms of brain recognition, which closely associates with other brain functions like perceiving and decision making ,and provides a framework and platform for the researches

J. Zhang (✉)
School of Information Science and Engineering, Institute for Cognitive Neurodynamics,
East China University of Science and Technology, Shanghai 200237, China
e-mail: ecustzjp@126.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_77,

of psychology, neurophysiology, artificial intelligence and image processing. It is widely believed that the brain contains specific mechanisms for visual motion processing [1, 2], and primary visual cortex (v1 area) constitutes the first stage for the detecting of local motion information.

The research on the classic receptive field with feedforward dominated afferent is extensive. However, the phenomenon of surrounding effect was raised up by physiological study, and the mechanism underlying this phenomenon is now an appealing subject to more and more researchers [3, 4], this interaction is hypothesized to be mediated by horizontal connectivities and feed back from higher cortex [5]. The precise function of this configuration is not clear, it is assumed the information is globally integrated and salient image features are "highlighted" in this manner [6]. This interaction seems to be related with low and mid-level tasks as contour linking and texture segmentation as well as higher level processes like perceptual organization, attention, and visual awareness [7]. Hansen and Neuman [8] proposed a recurrent model which enhanced the contour and suppressed the noise. The temporal interaction besides spatial configuration also drawing attention from researchers, e.g., Motoyoshi [9] investigated the temporal dynamics of mutual interactions between local detectors, and attribute this effect to the speed limitation of propagating activities, and proposed a model based on this idea. Peggy Series et.al also noticed the temporal property of mutual interaction between v1 neurons [10], and established a model on the basis of feedforward afferent and horizontal connectivities, the response under different model and stimulus configurations are studied.

We proposed a spatial-temporal integrating model v1 cortex model concerning directional selectivity of visual apparent motion in this article. This model employs two stages of processing: short-range competition and long-range integration. In the short-range competition, the max activated neuron suppresses the rest neurons in the neighborhood. As for long-range stage, the winner of the first stage modulates the neurons from a certain distance, facilitating the neurons with similar direction selectivity and inhibiting those with different preferred direction, while the temporal delay is concerned due to the finite speed of propagation of activity, which also induced the temporal integration of previous information.

2 Model Description

We proposed a plausible model for visual motion perception in primary visual cortex. This model includes local motion detectors constructed with feedforward afferent selective for specific speed and direction of moving stimulus, short-range inhibition that maxim activated neuron inhibited the rest in the neighborhood, and long-range interaction with respect to the distance, alignment and similarity, which also concerned the temporal delay due to finite propagating speed.

2.1 Feedforward Afferent

We can describe a moving scene in x-y-t space [11]. Thus we can use a spatial-temporal-direction selective unit to describe a local motion speed. Simoncelli and Heeger established a linear model which is actually an array of filters selective for different temporal, spatial frequencies and directions [12]

$$L_n(t) = \iiint s_n(x, y, T)A(x, y, t - T)dxdydT + \alpha \tag{1}$$

$$S_n(t) = \frac{k \lfloor L_n(t) \rfloor^2}{\sum_m \lfloor L_m(t) \rfloor^2 + \sigma^2} \tag{2}$$

Where $L_n(t)$ is the response of the nth simple cell, the weighting functions $s_n(x, y, t)$ are a set of directional third derivatives of Gaussian with different space-time orientations, $A(x, y, t)$ is the local contrast of the stimulus images. α represents the spontaneous firing rate, $\lfloor L(t) \rfloor^2 \equiv \max[0, L(t)]^2$ is the half-squaring operation. Equation (2) represents a normalization operation, where k is the maximum attainable response, and σ is the semi-saturation constant of normalization. V1 complex cell responses are computed as local averages of simple cell response:

$$C(t) = \sum_m c_{nm} s_m(t) \tag{3}$$

For the neuron at a certain point and time whose preferred direction θ was implied in the weighting functions $s_n(x, y, t)$, the response can be rewritten as $C(x, y, t, \theta)$.

2.2 Short-Range Competition

In this stage, the winner of the local cell suppress the responses of rest cells, and represents the motion information at corresponding points, for simplicity this operation can be descript as:

$$C(x, y, t) = \max[C(x, y, t, \theta)] \tag{4}$$

2.3 Long-Range Interaction

The localized neuron pools the signal from neurons in the relatively remote surrounding area, these interactions are correlate with the distances, the alignments and the differences of preferred direction between neurons. The net effect mediates the feedforward afferent, and eventually the response:

$$R(x, y, t, \theta) = C(x, y, t, \theta)\left(1 + \sum_{n \in D} C_n(t - \frac{r}{c})B(r, \varphi)\right) \tag{5}$$

Where D is the surrounding area, r is the distance, c is the propagating speed, $B(r, \varphi)$ represents the strength of the connectivity, which is a function of distance and the difference of preferred directions between neurons.

$$B(r, \varphi) = \eta^* \cos(\theta - \varphi)^* \exp(-r^2/2\sigma) \tag{6}$$

Thus the responses of neurons are not only spatially but also temporally integrated with respect to local configuration of the neuron alignments, in this manner, a contextual perception of moving visual stimulus is attained.

3 Results

We used different sequences of gray-scale images as model inputs, and the responses of different stimulus were derived. In our model, we employed a set of local motion energy detectors with different preferred directions to simulate v1 complex cells. The preferred directions of these detectors are distributed with an interval of 30°, and 12 of them share one location to cover 360°.

We first used a moving sin grating with spatial frequency of 0.05 cycles/pixel, temporal frequency of 0.05 cycles/frame. As shown in Fig. 1, the arrow indicates the moving direction, which is 45°.

Figure 2 shows the responses to sinusoidal gratings inputs. The left panel shows the responses of model with only feed-forward afferents, and the right panel is

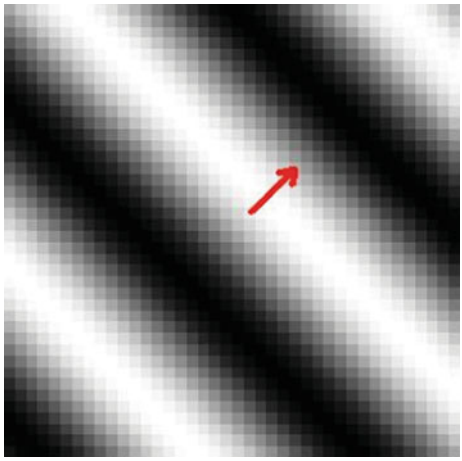


Fig. 1 A moving sinusoidal grating. with SF of 0.05 cycles/picel and TF of 0.05 cycles/frame. The *arrow* indicates the direction of motion

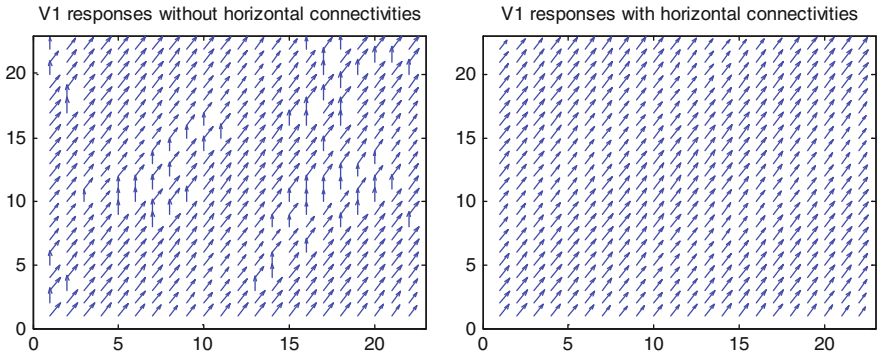


Fig. 2 The responses to sinusoidal grating input. The *left panel* is the responses of model with only feed-forward afferents, and the *right panel* is the responses with horizontal connectivities, the *arrows* indicate the preferred direction of most activated neurons

the responses of model with spatial-temporal integration due to horizontal interactions. For feed-forward model, most of the responses are in accordance with the input direction, with a portion deviate from the most likely direction. But for the model with horizontal interactions, because the right response dominates the site, those responses that don't agree with stimulus direction are inhibited due to spatial-temporal integration, which exhibits the ability of inhibition of noisy responses.

We also used a rectangular moving rightward as model input, which consists of 40 frames and moving at a speed of 1 pixel/frame, as shown in Fig. 3

While given input of moving rectangular, the neuron activations with different preferred direction exhibit significant disparity. As shown in Fig. 4, the neurons with preferred direction in accordance with the stimulus direction shows the maximum activation, especially at the two ends of rectangular, for the moving of luminance gradient in the left and right ends are most significant but none in the center of the rectangular. At the layer of neurons with preferred direction at 90°, the activation

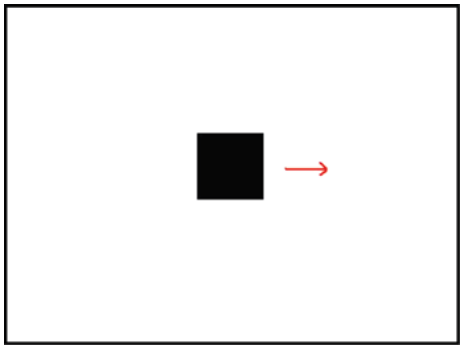


Fig. 3 Stimulus of a rectangular moving rightwards

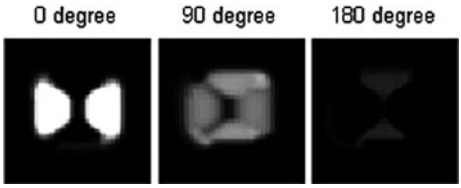


Fig. 4 Responses of v1 neurons with preferred direction of 0,90,180 (reverse from stimulus direction)deg respectively

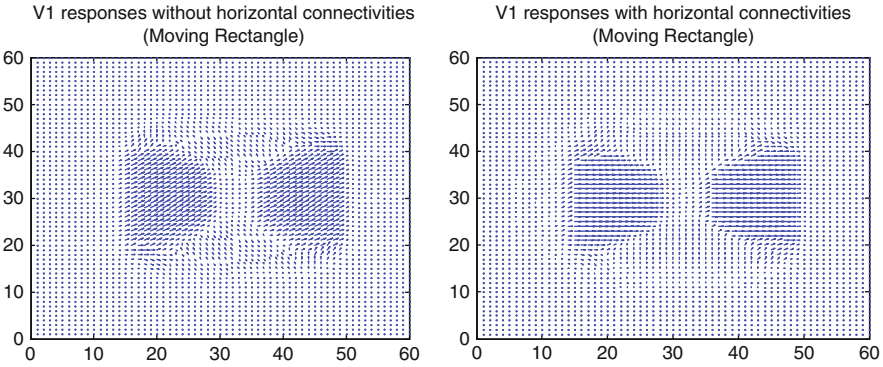


Fig. 5 The responses of model with only feed-forward afferent and model with horizontal connectivities

is less significant than that of 0°, and for the layer with opponent direction, the activation is totally inhibited by horizontal interactions.

For the model without horizontal interactions, as the left panel of Fig. 5 shows, the responding directions deviate from the input direction of stimulus image sequences, and intricately distributed over large direction range. But for model with horizontal interactions, the responding directions are consistent with the input direction, and concentrate in a small direction range. It can be explained like this: although the neurons with most alike directions with input don't necessary get the maximum activation, but they can most facilitated from surrounding neurons and previous "memory" via spatial and temporal integration, and contextual presentation of visual motion is attained in this manner.

4 Conclusion

The investigation of receptive field is a crucial component to reveal the mechanism of brain processing of visual information. Recently the surrounding mediating effect beyond classical RF is drawing researchers' attention, it is hypothesized that this mechanism involved in the global processing of visual information. An spatial-temporal integrating model concerning horizontal interactions was proposed to process the visual motion stimulus on the basis of the notion referred previously,

and good performances are attained in representing the directions of the moving images

Acknowledgement This work is supported by National Nature Science Foundation of China (10672057,10872068).

References

1. Maunsell, J.H.R., Newsome, W.T.: Visual processing in monkey extrastriate cortex. Annu. Rev. Neurosci. **10** (1987) 363401.
2. Albright, T.D.: Cortical processing of visual motion. In Miles, F.A., Wallman, J. eds.: Visual Motion and Its Role in the Stabilization of Gaze. Oxford: Elsevier (1993) 174–201.
3. Gilbert, C.D.: Horizontal integration and cortical dynamics. Neuron **9** (1992) 1–13.
4. Lamme, V.A.: The neurophysiology of figureground segregation in primary visual cortex. J. Neurosci. **15** (1995) 1605–1615.
5. Gilbert, C.D., Das, A., Ito, M., Kapadia, M., Westheimer, G. Spatial integration and cortical dynamics. Proc. Natl. Acad. Sci. USA **93** (1996) 615–622.
6. Lamme, V.A., Spekreijse, H.: Modulations of primary visual cortex activity representing attentive and conscious scene perception. Front. Biosci. **5** (2000) D232–D243.
7. Li, Z.: A neural model of contour integration in the primary visual cortex. Neural Comput. **10** (1998) 903–940.
8. Hansen, T., Neumann, H.: A recurrent model of contour integration in primary visual cortex. J. Vis. **8**(8) (2008) 8, 1–25.
9. Motoyoshi, I.: The role of spatial interactions in perceptual synchrony. J. Vis. **4** (2004) 352–361.
10. Seriès, P., Georges, S., Lorenceau, J., Frégnac, Y.: Orientation dependent modulation of apparent speed: A model based on the dynamics of feed-forward and horizontal connectivity in V1 cortex. Vis. Res. **42** (2002) 2781–2797.
11. Adelson, E.H., Bergen, J.R.: Spatiotemporal energy models for the perception of motion. Opt. Soc. Am. A **2**(2) (February 1985).
12. Simoncelli, E.P., Heeger, D.J.: A model of neuronal responses in visual area MT. Vis. Res. **38**(5) (1998) 743–761.

Periodic Motion Control of Human Arm Movement Based on CPG model

Jiong Wu, Rubin Wang, Zhikang Zhang, and Jianting Cao

Abstract In this paper, a Central Pattern Generator (CPG) model of a flexor and an extensor is developed, according to the experiments implemented in a crank-rotation task. The drive torque which achieved the arm movement is obtained by six muscles models consisted of outputs of the CPG model. The parameters of CPG model is optimized by minimum torque-change criterion and genetic algorithm. Finally, the simulation results are compared with to the experimental data and the CPG model of six muscles are discussed.

Keywords Central pattern generator · Periodic motion control · Six muscles models · Minimum torque change criterion · Genetic algorithm

1 Introduction

EMG signal is a potential change while the central nervous system planned muscle activity, which is related to the structure of muscles, mechanical properties of the movement and chemical changes during the muscle contraction. Therefore, changes in EMG signals are used to study how the central nervous system control and coordinate human arm movements. Since the results of signal identification could not only used as control signals of mechanical hand and the artificial limb, but also served as stimulation signals of neural prosthetic for the rehabilitation therapy of the human body, it is more and more important to identify the status of the human arm movement in the field of biomechanical research through the EMG signals.

In animal motions generation, a number of brain-nervous subsystems, such as cerebral cortex, cerebellum, basal ganglia, brain stem, spinal cord etc., are

J. Wu (✉)
School of Information Science and Engineering, Institute of Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, China
e-mail: iamwujiong@mail.ecust.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_78, © Springer Science+Business Media B.V. 2011

interdependently concerned, and the generated motions can be categorized into voluntary movement, automatic movement or reflex according to the concerned brain regions and the latent time. Among them, the automatic movement, such as walking, swimming, breathing etc. is a fundamental motion for life-sustaining, and has a periodic characteristic in common.

Based on the idea that trajectory planning must take arm dynamics into account, Uno et al. [1] suggested a minimum joint torque change criterion implying that the CNS plans the point-to-point reaching movements in the joint space based on the dynamic formulation. The criterion is given by

$$J = \int_0^T \dot{\boldsymbol{\tau}}^T \dot{\boldsymbol{\tau}} dt \tag{1}$$

Where τ is the combined vector of the joint torques.

Later on, Tadashi et al. [2] had confirmed that movement of human limbs is achieved by joint torques and each torque is specified as the sum of torques generated by muscle forces. As a result, it is verified that both a flexor and an extensor are activated throughout the entire movement and that the activation of muscles is controlled above a specific limit independent of the hand-held load. The minimum limit is reflected in the following criterion as a constraint of the neural input.

$$J = \frac{1}{2} \int_0^{t_f} \sum_{i=1}^{2} (u_i - f_i)^2 dt \tag{2}$$

The neural input is associated with muscle force by EMG signals and six-muscle model of human arm is developed firstly according to the anatomy theory which had played a major role in promoting the research of human arm movement.

In the previous work, it is proved that we can gain muscle force from EMG and the simulation results are pretty well matched to the experiment data [3]. The driving torque for motion control can be generated in proportion to the difference of the outputs of the extensor and flexor neurons [4]. For this reason, this paper proposed a CPG model based on the biarticular muscles of upper limb and built up a six-muscle model with other four muscle force function, then check this six-muscle model in Ken's dynamic model.

2 Rotate Crank Device and CPG Model

2.1 Rotate Crank Device

According to the crank experiment [5], it is supposed to compare the simulated results with the experiment results referred in ken's paper. (All the original data could be referred to in the paper mentioned above.)

Since the crank model had been well analyzed,

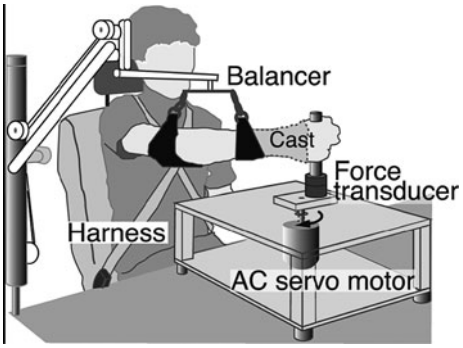


Fig. 1 The crank-rotation task

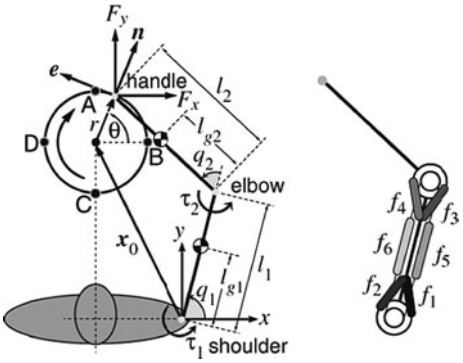


Fig. 2 Model of crank-rotation task. (**a**) Kinematic model of crank and arm. (**b**) Six muscle model

$$I\ddot{\theta} + B\dot{\theta} = re^T F, \tag{3}$$

$$M(q)\ddot{q} + h(q,\dot{q}) = \tau^* - J(q)^T F \tag{4}$$

$$\tau^* \equiv \tau - b\dot{q}, \tag{5}$$

$$J(q)\ddot{q} + \dot{J}(q,\dot{q})\dot{q} = r(\ddot{\theta}e - \dot{\theta}^2 n), \tag{6}$$

Where $F = \left[F_x, F_y\right]^T$ is the hand contact force, $\tau = [\tau_1, \tau_2]^T$ is the joint torque vector, e and n are the unit vectors of rotating contact frame, b stands for joint viscosity matrix.

2.2 CPG Model of Human Arm

The Matsuoka oscillator is the most frequently used, nonlinear oscillator model which constitute of CPG model. Two artificial neurons, connected by mutual inhibition, constitute the CPG model, and each neuron denotes flexor and extensor,

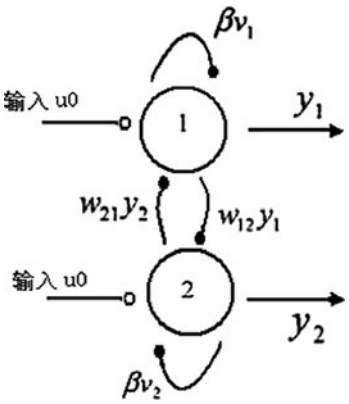


Fig. 3 CPG model of upper limb

respectively. The output of CPG is the difference of the outputs of the extensor and flexor neurons. The best feature of Matsuoka oscillator is that it could imitate the biological characteristics of organisms very well with or without external stimulus. As long as there is a stable constant stimulus, the output would be a stable oscillation wave. This output can be used as the control signal of muscle force, joint angle or driving torque.

According to the physiological structure of human arm, the biceps (extensor) and triceps (flexor) of upper limb mainly decide the rotation of elbow. Since it is so, the CPG of extensor and flexor is developed (Fig. 3) [6] and a muscle force vector is composed of the output and other four muscle function. Finally, the driving torque at the elbow is obtained.

$$\begin{aligned}
&\tau_1 \dot{u}_1 = -u_1 - \beta v_1 + \omega_{12} y_2 + u_0; \\
&\tau_2 \dot{v}_1 = -v_1 + y_1; \\
&\tau_1 \dot{u}_2 = -u_2 - \beta v_2 + \omega_{12} y_1 + u_0; \\
&\tau_2 \dot{v}_2 = -v_2 + y_2; \\
&y_i = f(u_i) = \max(u_i, 0), i = 1, 2
\end{aligned} \tag{7}$$

Here, the black bullets and hollow bullets denote inhibitory connections and excitatory connections, respectively. u_i and v_i correspond to internal state variables, membrane potential and membrane current. β is a fatigue constant, u_0 is a bias input, and ω_{ij} is a mutual inhibition weight between oscillators. For the sake of convenience, make $\omega_{12} = \omega_{21} = 0$ to simplify the mutual interaction of flexor and extensor. The rest parameters of CPG can be obtained through genetic algorithm. Take the rest parameters as initial sample of GA and minimum torque criterion as its Termination condition, program the whole model in the Mat Lab.

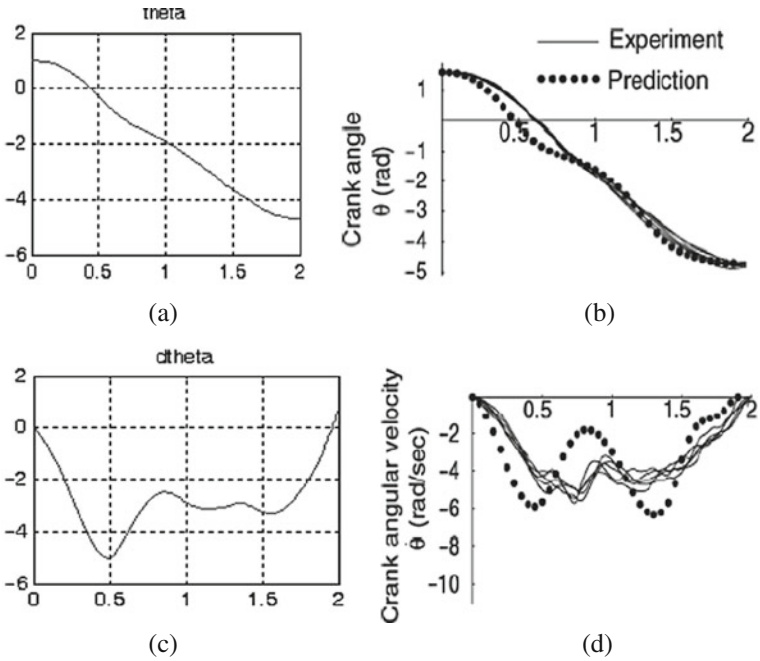


Fig. 4 (**a**, **c**) Obtained by CPG simulation. (**b**, **d**) Experimental data. *Solid lines* mean the experiment data and *dotted lines* mean the results from Ken et al. [5]

3 Comparison

Here, the parameters of CPG model are chosen to be the initial samples and minimum torque change criterion to be the termination conditions. The simulation results are as follows:

As we can see, the results obtained from CPG model show a great consistency and coherence than the simulation results in Ken's paper. It is also verified that the idea of applying CPG to period motion on the human arm is successful.

4 Conclusion

There are 7° of freedom in human upper limbs (excluding hands) and each one corresponds to the relative muscles. There would involve more muscles which many of them cannot be detected by a surface electrode, if the freedom of the human upper limb is not limited. In this paper, it would be just one degree of freedom at the elbow while the arm is restricted to make period motion in a plane. Since the elbow rotation of human arm movement mainly depend on biceps (flexor) and triceps (extensor), this paper has developed a CPG model of flexor and extensor of upper limb based on the neural oscillator of mutual inhibition and characteristics of extensor and flexor.

The simulation results are compared with to the experimental data and the CPG model of six muscles are discussed.

Acknowledgements This work is supported by National Nature Science Foundation of China (10672057, 10872068).

References

1. Uno, Y., Kawato, M., Suzuki, R.: Formation and control of optimal trajectory in human multijoint arm movement. Biol. Cybern. **61** (1989) 89–101.
2. Kashima, T., Isurugi, Y., Shima, M. Analysis of a muscular control system in human movements. Biol. Cybern. **82** (2000) 123–131.
3. Wu, J., Wang, R., Shen, E., Zhang, Z.: A new algorithm in the trajectory-calculating of human arm movement. J. Vib. Shock **27**(9) (2008) 147–149.
4. Kondo, T., Ito, K.: Periodic Motion Control by Modulating CPG Parameters Based on Time-Series Recognition. ECAL 2005, LNAI 3630 (2005) 906–915.
5. Ken, O., Mikhail, M.S., Zhiwei, L., Shigeyuki, H., Rafael L.: Optimal trajectory formation of constrained human arm reaching movements. Biol. Cybern. **91** (2004) 23–36.
6. Dong, W., Wang, R., Zhang, Z.: The simulation study on the pattern of muscles controlled by CPG in rhythm gait movement. J. Dyn. Control **6**(4) (2008) 327–331.

Part VIII
Global Cognitive Function

A Neurodynamical Approach to Visual Affordances: The Selective Attention for Action Model (SAAM)

Christoph Böhme and Dietmar Heinke

Abstract Classically, visual attention is assumed to be influenced by visual properties of objects. However, recent experimental evidence suggests that visual attention is also guided by action-related properties of objects ("affordances" [1]). In our model we use a neurodynamical approach to soft-constraint satisfaction. The model implements two sets of constraints: the first set accounts for anatomical constraints of hand postures; the second set considers suitable contact points on objects. We will demonstrate here that the model can successfully mimic human grasping and affordance-guided attentional behaviour.

Keywords Visual attention · Object affordance · Computational modelling · Human grasping

1 Introduction

Actions need to be tightly guided by vision in our daily interactions with our environment. To maintain such a direct guidance, Gibson postulated that the visual system automatically extract "affordances" of objects [1]. According to Gibson, affordances refer to parts or properties of visual objects that are directly linked to actions or motor performances. For instance, a handle of a cup *affords* directly a reaching and grasping action. Recently experimental studies have produced empirical evidence in support for this theory (e.g. [2–4]). Interestingly, recent experimental evidence also suggests that selective attention is guided towards action-relevant locations (e.g. [5]).

This paper aims to lay the foundations for a computational model of such affordance-based guidance of attention. We designed a neurodynamical model

C. Böhme (✉)
School of Psychology, University of Birmingham, Birmingham B15 2TT, UK
e-mail: cxb632@bham.ac.uk

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_79,

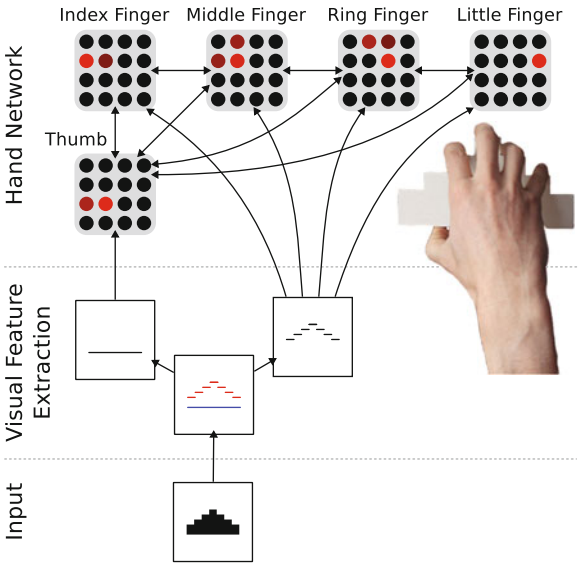


Fig. 1 Overview of the model structure

which determines contact points for a stable grasp of an object (see Fig. 1). The model extracts these contact points directly from the input image. Hence, such a model could be construed as an implementation of an automatic detection of object affordances for grasping. To realise the attentional guidance through affordances, we integrated the selection mechanisms employed in the Selective Attention for Identification Model (SAIM [6]). Since this new model performs selection for action rather than identification, we termed the new model Selective Attention for Action Model (SAAM).

2 The Selective Attention for Action Model (SAAM)

Figure 1 gives an overview of SAAM. The input consists of black&white images. The output of the model is generated in five "finger maps" of a "hand network". The finger maps encode the finger positions which are required for producing a stable grasp of the object in the input image. At the heart of SAAM's operation is the assumption that stable grasps are generated by taking into account two types of constraints: geometrical constraints imposed from the object shape and anatomical constraints given by the hand. In a soft-constraint satisfaction approach [7] constraints define activity patterns in the finger maps that are permissible and others that are not. An energy function was defined which generates minimal values by just these permissible activity values. To find these minima, a gradient descent procedure is applied resulting in a differential equation system. The differential equation system defines the topology of a biologically plausible network.

The geometrical constraints are extracted from the shape of the object in the visual feature extraction stage. To begin with, obviously, only edges constitute suitable contact points for grasps. Furthermore, edges have to be perpendicular to the direction of the forces exerted by the fingers. Hence only edges with a horizontal orientation make up good contact points, since we only consider a horizontal hand orientation in this first version of the model (see Fig. 1). To exert a stable grasp, thumb and fingers need to be located at opposing sides of an object. This requirement was realised by separating the edge-filtered input according to the direction of the gradient change at the edge to indicate appropriate locations for the fingers and the thumb (see Fig. 1). The results of the separation feed into the corresponding finger maps providing the hand network with the geometrical constraints.

The anatomical constraints implemented in the hand network take into account that the human hand cannot form every arbitrary finger configuration to perform grasps. For instance, the maximum grasp width is limited by the size of the hand and the arrangement of the fingers on the hand makes it impossible to place the index, middle, ring, and little finger in another order than this one. After applying the energy minimisation approach, these anatomical constraints are implemented by excitatory connections between the finger layers in the hand network (see Fig. 1). Each weight matrix defines how every single neuron of one finger map projects onto another finger map. Each neuron in the target map sums up all activation fed through the weight matrices. Finally, since a finger can be positioned at only one location, a winner-takes-all mechanism was implemented in all finger maps. In the simulation section we will show that this selection mechanism also implements global selection mimicking selective attention.

3 Studies: Single-Object Inputs and Two-Object Inputs

The first study tested whether SAAM can generate expedient grasps in general and whether these grasps mimic human grasps. To accomplish this, simulations with single objects in the visual field were compared with experimental data on grasping these objects.

Figure 2 shows an example of the simulation results. This illustration also include the mean finger positions from the experimental data for a comparison with the simulation data. The ellipses around the mean finger positions illustrate the variations

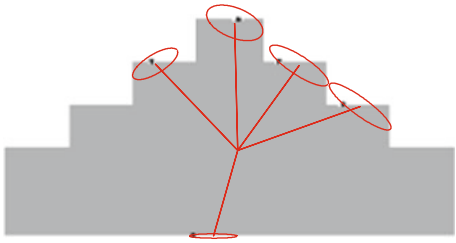


Fig. 2 Simulation and experimental data

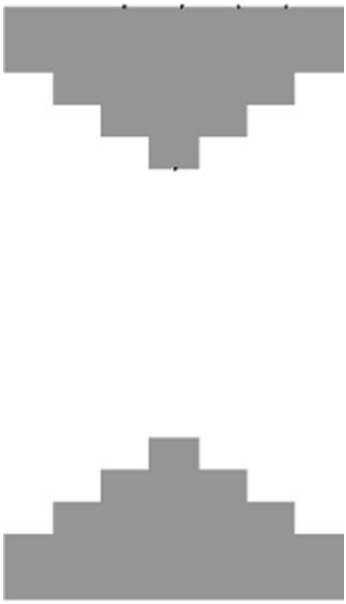


Fig. 3 Simulation of selective attention

in the data. The comparison shows that most finger positions lie within the ellipses. Hence the theoretical assumptions behind SAAM that geometrical and anatomical constraints are sufficient to mimic human behaviour have been confirmed.

The second set of simulations investigated SAAM's ability to simulate attentional processes by using input images with two objects. Figure 3 shows the simulation results. The simulations are successful in the sense that contact points for only one object were selected and the second object was ignored (see Conclusion for further discussions). Note that this is an emergent property of the interplay between all constraints. The geometrical and anatomical constraints ensure that only contact points around the object were selected and the WTA-mechanism restricts the contact points to one object. In addition, the weight matrices (anatomical constraints) determine the selection priorities of SAAM. At present we do not have reference data from humans. It would be especially interesting to see whether SAAM and humans have the same select preference.

4 Conclusion and Outlook

This paper set out to develop a model of affordance-based guidance of selective attention. To detect the parts of an object which afford a stable grasp, SAAM performs a soft-constraint satisfaction approach. The constraints were derived from the geometrical properties of the input object and the anatomical properties of the human hand. In a comparison between simulation results and experimental data

from human participants we could show that these constraints are sufficient to simulate human grasps. Note that an alternative approach would have been a complex torque and forces analysis. However, our simulations suggest that anatomical constraints render such an analysis obsolete. In a second set of simulations we tested whether SAAM cannot only extract object affordances but also implements the guidance of attention through affordances. Indeed, SAAM was able to select one of two objects based on their affordances. The interesting aspect here is that SAAM's performance is an emergent property from the interplay between the anatomical constraints. Especially, the competitive mechanism implemented in the finger maps is crucial for SAAM's attentional behaviour. This mechanism already proved important in the Selective Attention for Identification Model (SAIM [6]) for simulating attentional effects of human object recognition. Despite differences in the implementation of both models this similarity makes it conceivable that both models can be combined into one. In such a model SAIM's selection mechanism of whole objects can be guided by the SAAM's selection of contact points. Hence, this new model could integrate both mechanisms, selection by visual-properties and by action-related properties, forming a more complete model of selective attention.

Despite the successes reported here, this work is still in its early stages. First, we will need to verify the priorities of object selection predicted by SAAM in the second study. Also, there is a large amount of experimental data on the interaction between action knowledge and attention (see [8] for a summary). Therefore, we aim to integrate action knowledge into SAAM, e.g. grasping a knife for cutting or stabbing. With these extensions SAAM will sufficiently contribute to the understanding of how humans determine object affordances and how these lead to a guidance of attention.

References

1. Gibson, J.J.: The ecological approach to visual perception. Boston, MA: Houghton-Mifflin, (1979)
2. Grèzes, J., Decety, J.: Does visual perception of objects afford action? evidence from a neuroimaging study. Neuropsychologia **40**(2) (2002) 212–222.
3. Tucker, M., Ellis, R.: On the relations between seen objects and components of potential actions. J. Exp. Psychol. **24**(3) (1998) 830–846.
4. Borghi, A.M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., Nicoletti, R.: Are visual stimuli sufficient to evoke motor information? studies with hand primes. Neurosci. Lett. **411**(1) (2007) 17–21.
5. Handy, T.C., Grafton, S.T., Shroff, N.M., Ketay, S., Gazzaniga, M.S.: Graspable objects grab attention when the potential for action is recognized. Nat. Neurosci. **6**(4) (2003) 421–427.
6. Heinke, D., Humphreys, G.W.: Attention, spatial representation and visual neglect: Simulating emergent attention and spatial memory in the selective attention for identication model (SAIM). Psychol. Rev. **110**(1) (2003) 29–87.
7. Hopfield, J.J., Tank, D.W.: "Neural" computation of decisions in optimization problems. Biol. Cybern. **52**(3) (1985) 141–152.
8. Humphreys, G.W., Riddoch, M.J.: From vision to action and action to vision: a convergent route approach to vision, action, and attention. Psychol. Learn. Motiv. **42** (2003) 225–264.

A Comparative Analysis on Adaptive Modelling of Induced Feelings

Zulfiqar A. Memon, Jan Treur, and Muhammad Umair

Abstract Stimuli and activations of mental states usually induce emotional responses that are experienced by their associated feelings. This paper concentrates on how the strengths, by which such emotional responses are induced, depend on previous experiences. It presents a comparative analysis of three adaptive modelling approaches addressing how these induction strengths are adapted over time: Hebbian learning, temporal discounting and memory traces. Example simulation results are shown and commonalities and differences between the models are analysed.

Keywords Induced feelings · Adaptive · Memory traces · Temporal discounting · Hebbian learning

1 Introduction

For many responses to certain (external or internal) circumstances, an important role is played by experiences for similar circumstances in the past. How such circumstances are experienced does not only depend on the circumstances themselves but also on the extent to which emotional responses are induced and felt. For this paper it is assumed that such an induction process of experienced emotional responses takes the form of triggered preparations for body states that are in a recursive as-if body loop with certain feelings; e.g. [1–3]. The strengths by which stimuli or activations of mental states induce certain preparations for body states or (other) actions occurring as emotional responses, might be innate, but are often considered to be acquired, strengthened and/or adapted during lifetime; e.g. [4, 5]. These induction strengths of responses based on experiences from the past are assumed to play an

Z.A. Memon (✉)
Department of Artificial Intelligence, VU University Amsterdam, 1081 HV, Amsterdam, The Netherlands
e-mail: zamemon@few.vu.nl

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_80,

important role in a variety of behaviors, for example, involving decision making according to the Somatic Marker Hypothesis e.g. [4, 6].

To explain or model such a development of induction strength of a response, in the literature different perspectives can be found. In this paper three different alternatives for modelling approaches are considered. One of them is a *Hebbian learning* approach e.g. [7–9]. A second alternative is based on a *temporal discounting* approach as often is used in modelling intertemporal decision making or in modelling trust dynamics e.g. [10, 11]. The third alternative considered is a case-based memory modelling approach based on *memory traces* e.g. [12, 13]. Each of the three approaches is briefly presented and some simulation results for a common case study are discussed and compared to each other to find out under which circumstances the approaches coincide, and when they differ.

As a starting point, in Section 2 it is shown how elements from neurological theories on generation of emotion and feeling were adopted, and based on them first a computational model was set up that models the effect of stimuli on emotional responses and feelings. In accordance with such literature a converging recursive as-if body loop to generate a feeling is taken as a point of departure. Next, as the main step, in Section 3 the three different adaptation models for the strength of the emotional response were integrated in this model. In Section 4 simulation results are presented. Section 5 shows how the three resulting models were evaluated and compared based on simulation experiments. Finally, Section 6 is a discussion.

2 Dynamics of Emotional Responses and Feelings

As any mental state in a person, a sensory representation state induces emotions felt within this person, as described by [2, 6]; for example: "...Through either innate design or by learning, we react to most, perhaps all, objects with emotions, however weak, and subsequent feelings, however feeble." [6, p. 93]. In some more detail, emotion generation via an as-if body loop roughly proceeds according to the following causal chain; see Damasio [2, 6]: sensory representation of stimulus → preparation for body state → sensory representation of body state → feeling. The as-if body loop is extended to a recursive as-if body loop by assuming that the preparation of the bodily response is also affected by the state of feeling the emotion: feeling → preparation for body state as an additional causal relation. Such recursiveness is also assumed by Damasio [6], as he notices that what is felt by sensing is actually a preparation for a body state which is an internal object, under control of the person: "The object at the origin on the one hand, and the brain map of that object on the other, can influence each other in a sort of reverberative process that is not to be found, for example, in the perception of an external object." [6] p. 91. Thus the obtained model for emotion generation is based on reciprocal causation relations between emotion felt and preparations for body states. Within the model presented in this paper both the preparation for the bodily response and the feeling are assigned an (activation) level or gradation, expressed by a number. The

cycle is modelled as a positive feedback loop, triggered by a sensory representation and converging to a level of feeling and preparation for body state.

Informally described theories in scientific disciplines, for example, in biological or neurological contexts, often are formulated in terms of causal relationships or in terms of dynamical systems. To adequately formalise such a theory the hybrid dynamic modelling language LEADSTO has been developed that subsumes qualitative and quantitative causal relationships, and dynamical systems; cf. [14]. This language has been proven successful in a number of contexts, varying from biochemical processes that make up the dynamics of cell behaviour to neurological and cognitive processes e.g. [1, 15]. Within LEADSTO the *dynamic property* or temporal relation a $\rightarrow\!\!\!\!\rightarrow_D$ b denotes that when a state property a occurs, then after a certain time delay (which for each relation instance can be specified as any positive real number D), state property b will occur. Below, this D will be taken as the time step Δt, and usually not be mentioned explicitly. In LEADSTO both logical and numerical calculations can be specified in an integrated manner, and a dedicated software environment is available to support specification and simulation.

An overview of the basic model for the generation of emotional responses and feelings is depicted in Fig. 1. This picture also shows representations from the detailed specifications explained below. However, note that the precise numerical relations between the indicated variables V shown are not expressed in this picture, but in the detailed specifications below, labeled by LP1 to LP5 in the picture.

Note that the sensor and effector state for body states and the dashed arrows connecting them to internal states are not used in the model considered here. In the dynamic properties below capitals are used for variables (assumed universally quantified). First the part is presented that describes the basic mechanisms to generate a belief state and the associated feeling, starting with how the world state is sensed.

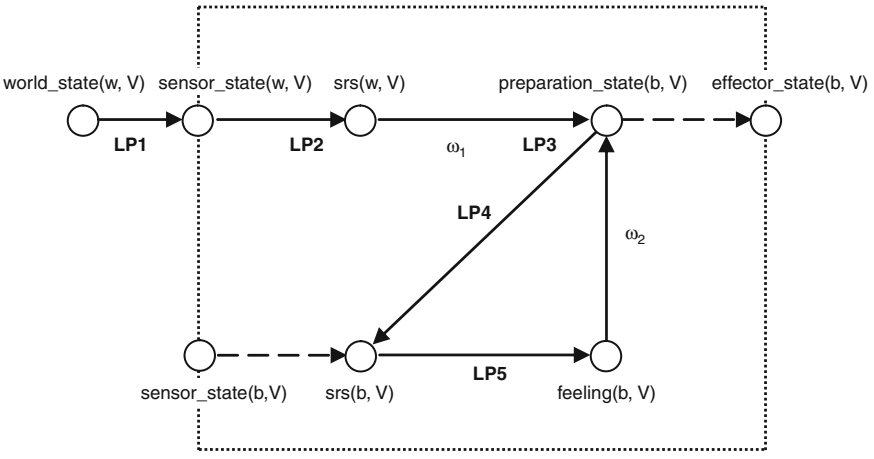


Fig. 1 Overview of the connections in the model for induced emotional responses and feelings

LP1 Sensing a world state
If world state property W occurs of strength V
then the sensor state for W will have strength V.
world_state(W, V) → sensor_state(W, V)

From the sensor states, sensory representations are generated according to the dynamic property LP2.

LP2 Generating a sensory representation for a sensed world state
If the sensor state for world state property W has strength V,
then the sensory representation for W with have strength V.
sensor_state(W, V) → srs(W, V)

Dynamic property LP3 describes the emotional response a sensory representation of a stimulus in the form of the preparation for a specific bodily reaction.

LP3 From sensory representation and feeling to preparation of a body state
If a sensory representation for w with strength V_1 occurs
and feeling the associated body state b has strength V_2
and the preparation state for b has strength V_3
and the connection from sensory representation of w to preparation for b has strength ω_1
and the connection from feeling to preparation for b has strength ω_2
and β is the person's orientation for emotional response
and γ is the person's flexibility for bodily responses
then after Δt the preparation state for body state b will have strength
$V_3 + \gamma(h(\beta, \omega_1, \omega_2, V_1, V_2) - V_3)\ \Delta t$.
srs(w, V_1) & feeling(b, V_2) & preparation_state(b, V_3) &
has_connection_strength(srs(w), preparation(b), ω_1) &
has_connection_strength(feeling(b), preparation(b), ω_2)
→→ preparation (b, $V_3 + \gamma$(h(β, ω_1, ω_2, V_1, V_2) − V_3) Δt)

The resulting level for the preparation is calculated based on a function $h(\beta, \omega_1, \omega_2, V_1, V_2)$ of the original levels. For the function $h(\beta, \omega_1, \omega_2, V_1, V_2)$ the following was taken:

$$h(\beta, \omega_1, \omega_2, V_1, V_2) = \beta(1 - (1 - \omega_1 V_1)(1 - \omega_2 V_2)) + (1 - \beta)\ \omega_1 \omega_2 V_1 V_2$$

Note that this formula describes a weighted sum of two cases. The most positive case considers the two source values as strengthening each other, thereby staying under 1: combining the imperfection rates $1 - \omega_1 V_1$ and $1 - \omega_2 V_2$ of them provides a decreased rate of imperfection expressed by $1 - (1 - \omega_1 V_1)(1 - \omega_2 V_2)$. The most negative case considers the two source values in a negative combination: combining the imperfections of them provides an increased imperfection. This is expressed by $\omega_1 \omega_2 V_1 V_2$. The parameter β can be used to model a characteristic that expresses the person's orientation for emotional response (from 0 as weakest response to 1 as strongest response). Dynamic properties LP4 and LP5 describe the as-if body loop.

LP4 From preparation to sensory representation of a body state
If preparation state for body state B occurs with strength V,
then the sensory representation for body state B with have strength V.
preparation(B, V) →→ srs(B, V)

LP5 From sensory representation of body state to feeling
If a sensory representation for body state B with strength V occurs,
then B will be felt with strength V.
srs(B, V) ← feeling(B, V)

3 Integrating Adaptation Models for the Induction Strengths

Three adaptation models for the induction strength ω_1 of the connection from sensory representation to preparation have been integrated. As a scenario it is assumed that over time different sensory representations occur in a repeated fashion. The first adaptation model presented follows a *Hebbian approach*. By this model the induction strength ω_1 of the connection from sensory representation to preparation is adapted using the following Hebbian learning rule. It takes into account a maximal connection strength 1, a learning rate η, and an extinction rate ζ.

LP6 Hebbian learning rule for connection from sensory representation of stimulus to preparation
If the connection from sensory representation of w to preparation of b has strength ω_1
and the sensory representation for w has strength V_1
and the preparation of b has strength V_2
and the learning rate from sensory representation of w to preparation of b is η
and the extinction rate from sensory representation of w to preparation of b is ζ
then after Δt the connection from sensory representation of w to preparation of b
will have strength $\omega_1 + (\eta V_1 V_2(1 - \omega_1) - \zeta\omega_1)\ \Delta t$.
has_connection_strength(srs(w), preparation(b), ω_1) & srs(w, V_1) & preparation(b, V_2) &
has_learning_rate(srs(w), preparation(b), η) & has_extinction_rate(srs(w), preparation(b), ζ)
→ has_connection_strength(b, w, $\omega_1 + (\eta V_1 V_2(1 - \omega_1) - \zeta\omega_1)\ \Delta t$)

A similar Hebbian learning rule can be found in [8, p. 406]. As a next model a *temporal discounting principle* is used to adapt the induction strength ω_1 of the connection from sensory representation to preparation.

LP7a Temporal discounting learning rule for sensory representation of stimulus
If the connection from sensory representation of w to preparation of b has strength ω_1
and the sensory representation for w has strength V and $V>0$
and the discounting rate from sensory representation of w to preparation of b is α
and the extinction rate from sensory representation of w to preparation of b is ζ
then after Δt the connection from sensory representation of w to preparation of b
will have strength $\omega_1 + (\alpha(V - \omega_1) - \zeta\omega_1)\ \Delta t$.
has_connection_strength(srs(w), preparation(b), ω_1) & srs(w, V) & V>0 &
has_discounting_rate(srs(w), preparation(b), α) & has_extinction_rate(srs(w), preparation(b), ζ)
→ has_connection_strength(srs(w), preparation(b), $\omega_1 + (\alpha(V - \omega_1) - \zeta\omega_1)\ \Delta t$)

LP7b Temporal discounting learning rule for sensory representation of stimulus
If the connection from sensory representation of w to preparation of b has strength ω_1
and the sensory representation for w has strength 0
and the extinction rate from sensory representation of w to preparation of b is ζ
then after Δt the connection from sensory representation of w to preparation of b
will have strength $\omega_1(\alpha(V - \omega_1) - \zeta\omega_1)\ \Delta t$.
has_connection_strength(srs(w), preparation(b), ω_1) & srs(w, 0) &
has_extinction_rate(srs(w), preparation(b), ζ)
→ has_connection_strength(srs(w), preparation(b), $\omega_1 - \zeta\omega_1\ \Delta t$)

The third model integrated is based on *memory traces*. Suppose is_followed_by(γ, a, b) indicates that within memory trace with identification label γ state a is followed by state b. The states addressed are the sensory representation state and the subsequent preparation state. It is assumed that each new pair of events <sensory representation, preparation> gets a new unique identification label γ, for example, based on a time stamp. The idea is then that for given states a and b, the strength of the induction from a to b is extracted in a case-based manner by the fraction of all represented traces in which a occurs with b as a next state from all traces in which a occurs:

$$\#\{\gamma \mid \text{is\_followed\_by}(\gamma, a, b)\} / \#\{\gamma \mid \exists c \text{ is\_followed\_by}(\gamma, a, c)\}$$

A temporal element can be incorporated by giving more weight to more recently represented memory traces. This was modelled by using temporal discounting when extracting the induction strength from the represented memory traces. Moreover, also levels of activations of both states were taken into account. In this approach traces get weights depending on their time label where traces that occurred further back in time have lower weights. The following rules represent how information is extracted from the time-labeled representations by counting the discounted numbers of occurrences.

LP8 Discounting memory traces
If the sensory representation for w has strength V_1
and the preparation of b has strength V_2
and the discounted number of memory traces with state srs(w) is X
and the discounted number of memory traces with state srs(w) and successor state preparation(b) is Y
and the discounting rate from sensory representation of w to preparation of b is α
then the discounted number of memory traces with state srs(w) is $\alpha X + (1 - \alpha)\ V_1$
and the discounted number of memory traces with state srs(w) and successor state preparation(b) is $\alpha Y + (1 - \alpha)V_1 V_2$
srs(w, V_1) & preparation(b, V_2) & has_discounting_rate(srs(w), preparation(b), α) &
memory_traces_including(srs(w), X) & memory_traces_including_both(srs(w), preparation(b), Y)
→→ memory_traces_including(srs(w), $\alpha X + (1-\alpha)\ V_1$) &
memory_traces_including_both(srs(w), preparation(b), $\alpha Y + (1-\alpha)V_1 V_2$)

Given these numbers the induction strength of the connection from sensory representation to preparation state is determined as Y/X.

LP9 Generation of preparations based on discounted memory traces
zzIf the discounted number of memory traces with state srs(w) is X
and the discounted number of memory traces with state srs(w) and successor state preparation(b) is Y
then the connection strength from srs(w) to preparation(b) is Y/X
memory_traces_including(srs(w), X) & memory_traces_including_both(srs(w), preparation(b), Y)
→→ has_connection_strength(srs(w), preparation(b), Y/X)

4 Example Simulation Results

Based on the computational model described in the previous section, a number of simulations have been performed. Some example simulation traces were included in this section as an illustration; see Figs. 2, 3 and 4, for the Hebbian learning, temporal

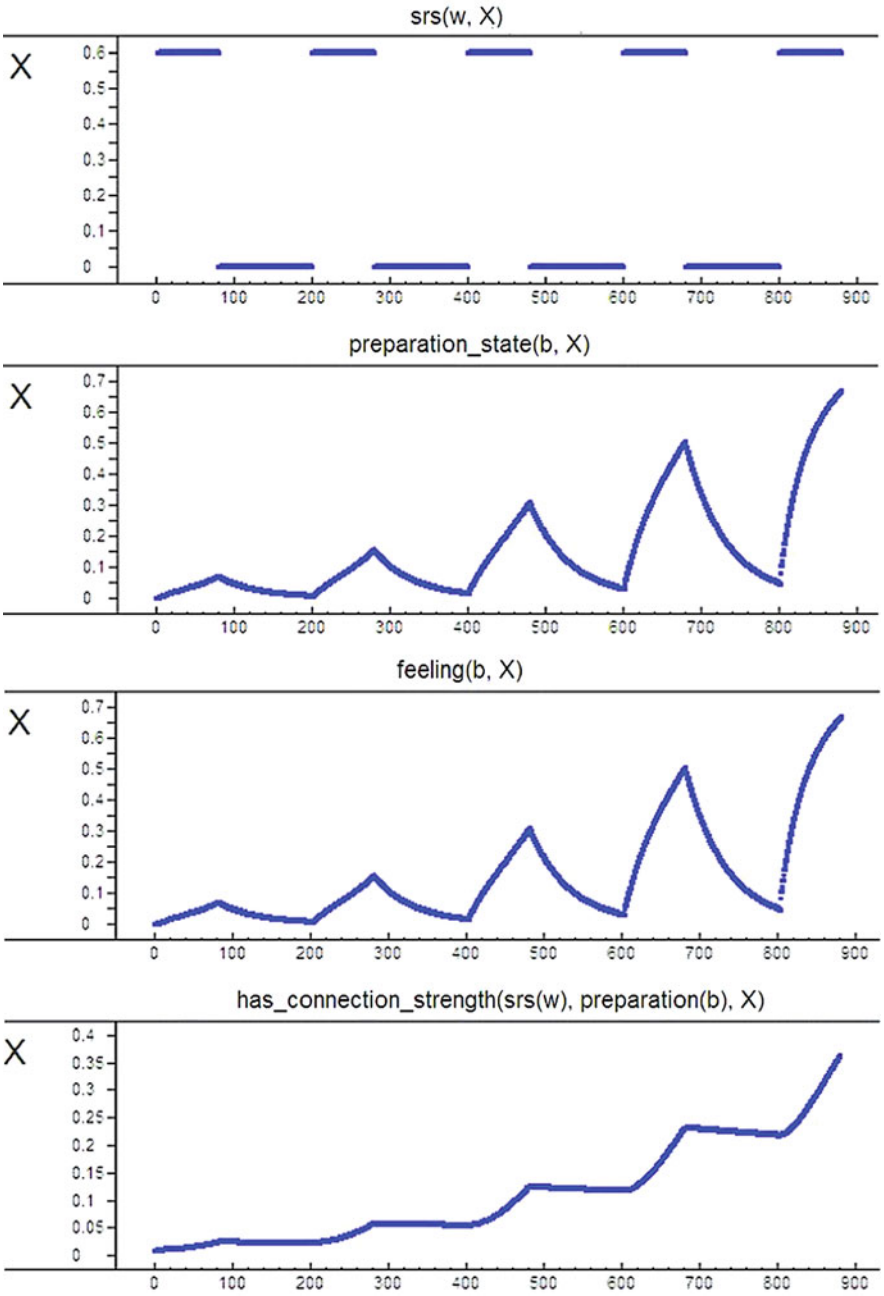


Fig. 2 Adaptation by hebbian learning (V1=0.6, β=0.9, γ=0.3, η=0.01, ζ=0.0005)

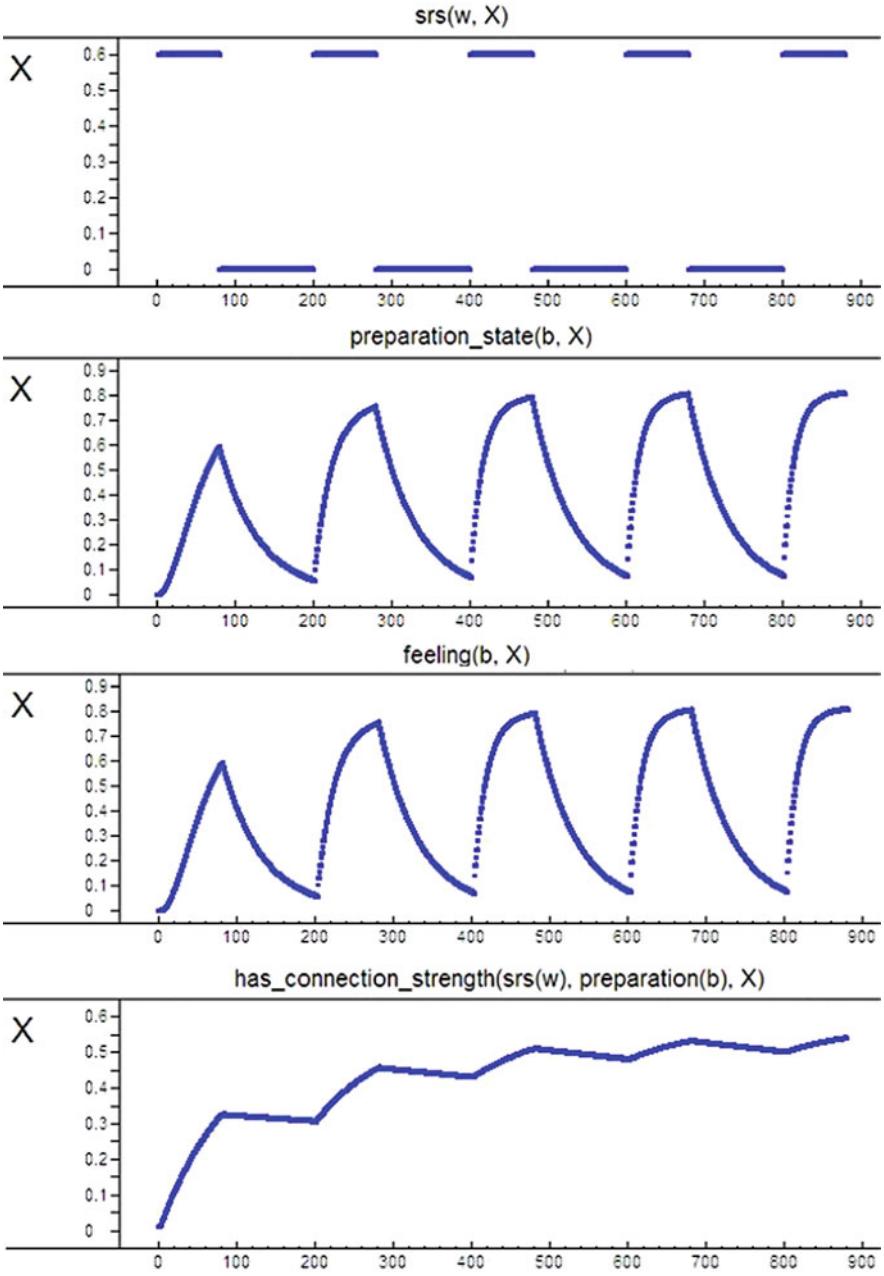


Fig. 3 Adaptation by temporal discounting (V=0.6, β=0.9, γ=0.3, α=0.01, ζ=0.0005)

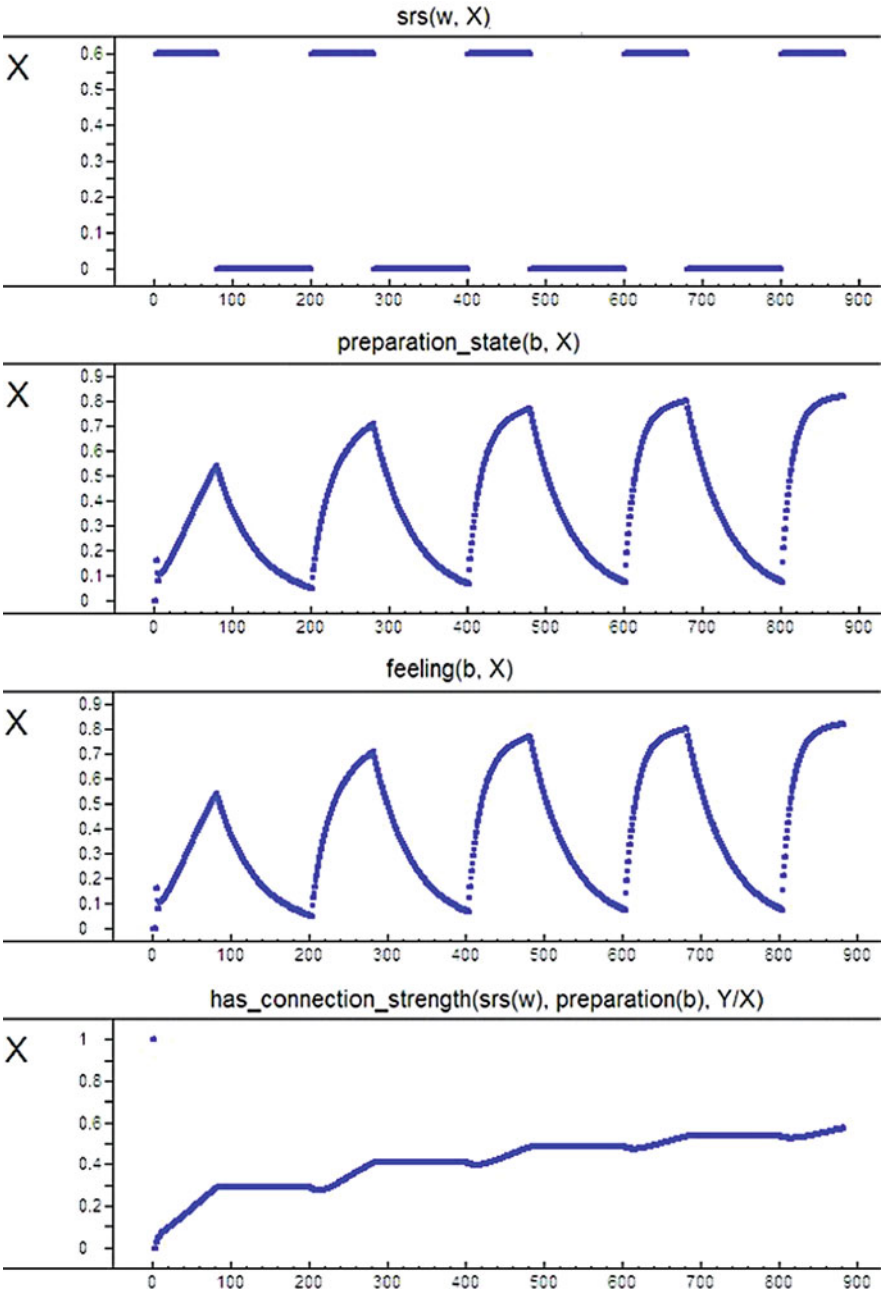


Fig. 4 Adaptation by memory traces (V1=0.6, β=0.9, γ=0.3, α=0.5, ζ=0.0005)

discounting and memory traces approach, respectively (here the time delays within the temporal LEADSTO relations were taken 1 time unit). Note that only a selection of the relevant nodes (represented as state properties in LEADSTO) is shown. In all of these figures, where time is on the horizontal axis, and the activation levels of the different state properties are on the vertical axis, quantitative information for state properties values for the different time periods are shown (by the dark lines). The activation levels of the state properties gradually increase while the sensory representation of the stimulus occurs, following the recursive feedback loop discussed in Section 2. These levels sharply decrease after the sensory representation of the stimulus stops occurring, as described by the temporal relationship LP3 in Section 2. Moreover, except for some decrease due to extinction, the induction strength of the connection from sensory representation to preparation state keeps its value in the phases without stimulus, until the sensory representation of the stimulus (again) occurs, as described by temporal relationship LP6 in case of Hebbian learning, LP7a and LP7b in case of temporal discounting and by LP8a and LP8b in case of memory traces. Further comparison of the three adaptation models are discussed in Section 5.

Figure 2 shows the adaptation model following the Hebbian approach. As can be seen in this figure, for sensory representation activation level 0.6 and initial level of preparation state 0, during the phases of the stimulus the activation levels of the preparation and feeling states progressively increase over time until they reach levels close to 0.9.

The induction strength which initially was set to 0.01, gradually increases to attain a strength around 0.8. The occurrence of this pattern is in line with the mathematical analysis which is discussed in the next section.

Figure 3 shows the temporal discounting approach. For sensory representation activation level 0.6 and initial level of preparation state 0, the activation levels of the preparation and feeling states gradually increase over the time until they reach values close to 0.8. The induction strength state initially set at 0.01, gradually increases to attain a strength around 0.55. Also the occurrence of this pattern is in line with the mathematical analysis discussed in the next section.

Figure 4 shows the adaptation model following the memory traces approach. As can be seen in the figure, for the sensory representation activation level 0.6 and initial level of preparation 0, the activation levels of preparation and feeling states gradually increases over time until levels close to 0.9 are reached. The induction strength gradually increases from initial value 0.01 to values around 0.7. The occurrence of this pattern also is in line with results from the next section.

5 Comparative Analysis of the Three Adaptation Models

This section compares the results of simulation for the three adaptive dynamic modeling approaches and presents some of the results of a mathematical analysis of the model that has been undertaken.

5.1 Comparison of Simulation Results

For a brief overview of a comparison of simulation results, see Table 1. The adaptation speed of the induction strength in the temporal discounting approach is faster as compared to other two approaches (see Figs. 2–4), even though the discounting rate for memory traces approach was set at a higher value. Notice also that the pattern of the learning curves differs: Hebbian learning shows a slow start, but later on gets more speed, while the other cases show a more or less opposite pattern. Moreover, the activation levels of the preparation and feeling states in the temporal discounting and memory traces approach increase faster as compared to the Hebbian approach.

The memory traces approach persist the strength more as compared to the temporal discounting approach, and the temporal discounting approach persist the strength more as compared to the Hebbian approach.

Below, some of the results of a mathematical analysis of possible equilibria of the model that has been undertaken are discussed. For an overview see also Table 2. Note that an equilibrium of the model involves constant values both for activation levels and connection strengths and it is also assumed that the stimulus is constant. Moreover, to avoid too many exceptional cases, it is assumed that the values for parameters γ, η,ζ are nonzero.

Table 1 Overview of outcomes of the example simulations for the three approaches

	Hebbian learning	Temporal discounting	Memory traces
Maximal induction strength reached	0.8	0.55	0.7
Adaptation speed	Lowest	Highest	Middle
Adaptation pattern	Slow start – fast finish	Fast start – slow finish	Fast start – slow finish
Extinction speed	Highest	Middle	Lowest
Maximal preparation and feeling levels	0.9	0.8	0.9
Speed in preparation and feeling levels	Lowest	Highest	Middle

Table 2 Overview of expressions for a number of possible equilibria

	General case	Case $V_1 = 1$ & $\omega_2 = 1$
Hebbian Learning	$V_2 = \beta\omega_1 V_1/(1 - \beta\omega_2 + (2\beta - 1)\omega_1\omega_2 V_1)$ $\omega_1 = \eta V_1 V_2/(\zeta + \eta V_1 V_2$ $= 1/(\zeta/\eta V_1 V_2) + 1)$	$\omega_1 = (((\eta_-(1/\beta - 1)\zeta)/$ $(2\zeta + (\eta_-)(1/\beta)\zeta))$ $V_2 = 1_-(1 - \beta)\zeta/\beta\eta$ $= 1_-(1/\beta - 1)\zeta/\eta$
Temporal Discounting	$V_2 = \beta\omega_1 V_1/(1 - \beta\omega_2 + (2\beta - 1)\omega_1\omega_2 V_1)$ $\omega_1 = V_1/(\zeta/\alpha + 1) \leq V_1$	$\omega_1 = 1/(\zeta/\alpha + 1)$ $V_2 = 1/(\zeta/\alpha(1/\beta - 1) + 1)$
Memory Traces	$\omega_1 = V_2 = (\beta V_1 - 1 + \beta\omega_2\_/(2\beta - 1)\omega_2 V_1$	$\omega_1 = V_2 = 1$

5.2 Equilibrium for Activation Level of Preparation State

First the equilibrium for the activation level of the preparation state has been investigated, expressed in the following relation derived from LP3 (note that due to LP4 the feeling level V_2 and preparation level V_3 are equal in an equilibrium; see):

$$\gamma(\beta(1-(1-\omega_1 V_1)(1-\omega_2 V_2))+(1-\beta)\,\omega_1\omega_2 V_1 V_2 - V_2) = 0$$

Assuming $\gamma \neq 0$, this equation can be solved by expressing V_2 into the other variables among which V_1 that denotes the activation level of the sensory representation.

$$\begin{array}{l}
\beta(1-(1-\omega_1 V_1)(1-\omega_2 V_2))+(1-\beta)\omega_1\omega_2 V_1 V_{2-} V_2 = 0 \Leftrightarrow \\
\beta(\omega_1 V_1+\omega_2 V_2-\omega_1\omega_2 V_1 V_2) = V_2-\omega_1\omega_2 V_1 V_2+\beta\ \omega_1\omega_2 V_1 V_2 \Leftrightarrow \\
\beta\omega_1 V_1 = V_2-\omega_1\omega_2 V_1 V_2+\beta\ \omega_1\omega_2 V_1 V_{2-}\ \beta\omega_2 V_2+\beta\omega_1\omega_2 V_1 V_2 \Leftrightarrow \\
\beta\omega_1 V_1 = V_2-\beta\omega_2 V_2+(2\beta-1)\omega_1\omega_2 V_1 V_2 \Leftrightarrow \\
\beta\omega_1 V_1 = (1-\beta\omega_2+(2\beta-1)\omega_1\omega_2 V_1)V_2 \Leftrightarrow \\
V_2 = \beta\omega_1 V_1/(1-\beta\omega_2+(2\beta-1)\omega_1\omega_2 V_1)
\end{array} \quad (1)$$

For 3 example values of $\beta\omega_1 V_1 = (1-\beta\omega_2+(2\beta-1)\omega_1\omega_2 V_1)V_2$ reduces to

$$\begin{array}{lll}
\beta = 0 & 0 = (1-\omega_1\omega_2 V_1)V_2 & \Leftrightarrow V_2 = 0 \quad \text{OR} \quad \omega_1 = \omega_2 = V_1 = 1 \\
\beta = 0.5 & 0.5\ \omega_1 V_1 = (1-0.5\ \omega_2)V_2 & \Leftrightarrow \omega_1 V_1 = (2-\omega_2)V_2 \Leftrightarrow \\
 & & V_2 = \omega_1 V_1/(2-\omega_2) \\
\beta = 1 & \omega_1 V_1 = (1-\omega_2+\omega_1\omega_2 V_1)V_2 & \Leftrightarrow V_2 = \omega_1 V_1/(1-\omega_2+\omega_1\omega_2 V_1)
\end{array}$$

For $V_1 = 1$ Eq. (1) is reduced to

$$V_2 = \beta\omega_1/(1-\beta\omega_2+(2\beta-1)\omega_1\omega_2)$$

For $\omega_2 = 1$ this is

$$V_2 = \beta\omega_1 V_1/(1-\beta+(2\beta-1)\omega_1 V_1)$$

For $\omega_2 = 1$ this is for the three values of β

$$\begin{array}{ll}
\beta = 0 & V_2 = 0 \quad \text{OR} \quad \omega_1 = \omega_2 = V_1 = 1 \\
\beta = 0.5 & V_2 = \omega_1 V_1 \\
\beta = 1 & V_2 = \omega_1 V_1/(\omega_1 V_1) = 1
\end{array}$$

For both $V_1 = 1$ and $\omega_2 = 1$ the equation is

$$V_2 = \beta\omega_1/(1-\beta+(2\beta-1)\omega_1) \Leftrightarrow V_2 = \beta/((1-\beta)/\omega_1+2\beta-1) \quad (2)$$

5.3 Hebbian Approach

Next the equilibrium for the connection strength ω_1 from sensory representation to preparation is analyzed for the Hebbian approach; this is expressed in the following relation derived from LP6:

$$\eta V_1 V_2(1-\omega_1) - \zeta\omega_1 = 0$$

For cases that V_1 and V_2 are nonzero, this can be used to express ω_1 as follows

$$\begin{aligned}&\eta V_1 V_2 = (\zeta + \eta V_1 V_2)\omega_1 \Leftrightarrow \omega_1 = \eta V_1 V_2/(\zeta + \eta V_1 V_2) \Leftrightarrow \\ &\omega_1 = 1/(\zeta(\eta V_1 V_2) + 1)\end{aligned} \tag{3}$$

In principle the two Eqs. (1) and (3) in ω_1 and V_2 can be explicitly solved, but for the general case this provides rather complex expressions for ω_1 and V_2. Therefore only the specific case $V_1 = 1$ and $\omega_2 = 1$ is pursued further.

Case $V_1 = 1$ and $\omega_2 = 1$ for the Hebbian approach

For $V_1 = 1$ Eq. (3) can be rewritten into

$$\omega_1 = 1/(\zeta(\eta V_2) + 1) \Leftrightarrow 1/\omega_1 = (\zeta/\eta V_2) + 1)$$

Substituting the above equation in the expression (2) for V_2 provides:

$$\begin{aligned}&V_2 = \beta/((1-\beta)/\omega_1 + 2\beta - 1) \Leftrightarrow ((1-\beta)/\omega_1 + 2\beta - 1)V_2 = \beta \Leftrightarrow \\ &((1-\beta)(\zeta/(\eta V_2) + 1) + 2\beta - 1)V_2 = \beta \Leftrightarrow \\ &(1-\beta)(\zeta/\eta + V_2) + 2\beta\ V_2 - V_2 = \beta \Leftrightarrow \\ &(\zeta/\eta + V_2) - \beta(\zeta/\eta + V_2) + 2\beta\ V_2 - V_2 = \beta \Leftrightarrow \\ &\zeta/\eta - \beta(\zeta/\eta + V_2) + 2\beta\ V_2 = \beta \Leftrightarrow \\ &\zeta/\eta - \ \beta\ \zeta/\eta + \beta V_2 = \beta \Leftrightarrow \beta\ V_2 = \beta \Leftrightarrow \beta\ V_2 = \beta\_\zeta/\eta + \beta\ \zeta/\eta \Leftrightarrow \\ &V_2 = 1\_(1-\beta)\xi/\beta\eta = 1\_(1/\beta - 1)\zeta/\eta\end{aligned}$$

From this an expression for ω_1 can be determined:

$$\begin{aligned}&\omega_1 = 1/(\zeta(\eta V_2) + 1) \Leftrightarrow \omega_1 = 1/(\zeta(\eta(1\_(1/\beta - 1)\zeta/\eta + 1) \Leftrightarrow \\ &\omega_1 = 1/(\zeta((\eta\_(1/\beta - 1)\zeta) + 1) \Leftrightarrow \omega_1 = (\eta\_(1/\beta - 1)\zeta)/(\zeta + (\eta\_(1/\beta - 1)\zeta)) \\ &\Leftrightarrow \\ &\omega_1 = (\eta\_(1/\beta - 1)\zeta)/(2\zeta + (\eta\_(1/\beta)\zeta))\end{aligned}$$

For the three example values of β the equation $\beta\ V_2 = \beta\_\zeta/\eta + \beta\ \zeta/\eta$ reduces to

$$\begin{aligned}&\beta = 0 \text{ This is impossible for nonzero } \zeta,\ V_1 \text{ and } V_2 \\ &\beta = 0.5 \quad 0.5\ V_2 = 0.5\_\zeta/\eta + 0.5\ \zeta/\eta \\ &\qquad\quad V_2 = 1\_\zeta/\eta \\ &\beta = 1 \qquad V_2 = 1\_\zeta/\eta\ \zeta/\eta = 1\end{aligned}$$

5.4 Temporal Discounting Approach

For the temporal discounting approach the variable V_2 does not play a role in the adaptation method. Temporal relation LP7a implies that for an equilibrium it holds:

$$\begin{aligned} &\alpha(V_1 - \omega_1) - \zeta\omega_1 = 0 \Leftrightarrow \alpha(V_1 - \omega_1) = \zeta\omega_1 \Leftrightarrow \\ &\alpha V_1 = \zeta\omega_1 + \alpha\omega_1 = (\zeta + \alpha)\,\omega_1 \Leftrightarrow \\ &\omega_1 = \alpha V_1/(\zeta + \alpha) = V_1/(\zeta/\alpha + 1) \end{aligned} \tag{4}$$

Note that since $(\zeta/\alpha + 1) \geq 1$ it follows that always $\omega_1 \leq V_1$ which indeed was observed in the simulations.

Case $V_1 = 1$ and $\omega_2 = 1$ for the temporal discounting approach

For $V_1 = 1$ expression (4) becomes $\omega_1 = 1(\zeta/\alpha + 1)$. By Eq. (2) for V_2 it follows

$V_2 = \beta/((1-\beta)/\omega_1 + 2\beta - 1) \Leftrightarrow V_2 = \beta/((1-\beta)\,(\zeta/\alpha + 1) + 2\beta - 1) \Leftrightarrow$
$V_2 = \beta/((\zeta/\alpha + 1)$ - $\beta(\zeta/\alpha + 1) + 2\beta - 1) \Leftrightarrow$
$V_2 = \beta/(\zeta/\alpha - \beta(\zeta/\alpha + 1) + 2\beta) \Leftrightarrow$
$V_2 = \beta/(\zeta/\alpha - \beta(\zeta/\alpha) + \beta) \Leftrightarrow V_2 = \beta/(\zeta/\alpha(1-\beta) + \beta) \Leftrightarrow$
$V_2 = 1/(\zeta/\alpha(1/\beta - 1) + 1)$

For the three example values of β the equation $V_2 = \beta/(\zeta/\alpha(1-\beta) + \beta)$ reduces to

$$\begin{aligned} &\beta = 0 \quad V_2 = 0 \\ &\beta = 0.5 \quad V_2 = 0.5/(\zeta/\alpha\ 0.5 + 0.5) = 1(\zeta/\alpha + 1) \\ &\beta = 1 \quad V_2 = 1 \end{aligned}$$

5.5 Memory Traces Approach

For the memory traces approach in an equilibrium the expression $\omega_1 = Y/X$ should remain the same, although in principle X and Y still may change. So the criterion is

$$Y + \Delta Y/X + \Delta X = Y/X$$

which can be rewritten as

$(Y + \Delta Y)X = (X + \Delta X)Y \Leftrightarrow YX + \Delta Y\,X = XY + \Delta X\,Y \Leftrightarrow X\,\Delta Y = Y\,\Delta X \Leftrightarrow$
$\Delta Y/\Delta X = Y/X$

Therefore according to temporal relation LP8a for an equilibrium it holds:

$$(\alpha\ V_1\ V_2 - \zeta Y)/(\alpha\ V_1 - \zeta X) = \mathrm{Y/X}$$

This can be rewritten as:

$(a\ V_1\ V_2 - \zeta Y)X = (\alpha\ V_1 - \zeta X)\mathrm{Y} \Leftrightarrow \alpha\ V_1\ V_2 X\zeta YX = \alpha V_1 Y - \zeta XY \Leftrightarrow$
$\alpha\ V_1\ V_2 X = \alpha\ V_1\ Y \Leftrightarrow$
$V_2 X = Y \quad \mathrm{OR} \quad V_1 = 0 \Leftrightarrow$
$\omega_1 = Y/X = V_2\ \mathrm{OR}\ \ V_1 = 0$

This can be used to obtain a value for both ω_1 and V_2 as follows:

$$
\begin{array}{ll}
\beta\omega_1 V_1 = (1 - \beta\omega_2 + (2\beta - 1)\omega_1\omega_2 V_1)V_2 \Leftrightarrow & \\
\beta\ V_2 V_1 = (1 - \beta\omega_2 + (2\beta - 1)\ V_2\omega_2 V_1)V_2 \Leftrightarrow & \\
\beta\ V_1 = 1 - \beta\omega_2 + (2\beta - 1)\ V_2\omega_2 V_1 & \text{OR } V_2 = 0 \Leftrightarrow \\
\beta\ V_1 = 1 + \beta\omega_2 = (2\beta - 1)\ V_2\omega_2 V_1 & \text{OR } V_2 = 0 \Leftrightarrow \\
\omega_1 = V_2 = (\beta V_1 - 1 + \beta\omega_2)/2(\beta - 1)\omega_2 V_2 & \text{OR } V_2 = 0 \text{ OR} \\
 & \beta = 0.5\ \&\ V_1 = 2 - \omega_2 \Leftrightarrow \\
\omega_1 = V_2 = (\beta\ V_1 - 1 + \beta\omega_2)/(2\beta - 1)\omega_2 V_1 & \text{OR } V_2 = 0 \text{ OR} \\
 & \beta = 0.5 \& V_1 = \omega_2 = 1
\end{array}
$$

For the three specific example values of β the following is obtained.

$$
\begin{array}{lll}
\beta = 0 & \omega_1 = V_2 = 0 & \text{OR } \omega_1 = \omega_2 = V_1 = V_2 = 1 \\
\beta = 0.5 & \omega_1 = V_2 = \omega_1 V_1/(2 - \omega_2) \Leftrightarrow & \\
 & \omega_1 = V_2 = 0 & \text{OR } 1 = V_1/(2 - \omega_2) \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = 2 - \omega_2 \quad \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = \omega_2 = 1\ \omega_1 = V_2 \\
\beta = 1 & \omega_1 = V_2 = \omega_1 V_1/(1 - \omega_2 + \omega_1\omega_2 V_1) & \quad \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } 1 - V_1(1 - \omega_2 + \omega_1\omega_2 V_1) \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = 1 - \omega_2 + \omega_1\omega_2 V_1 \quad \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } (1 - \omega_1\omega_2)V_1 = 1 - \omega_2 \quad \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = (1 - \omega_2)/(1 - \omega_1\omega_2)\&\omega_1 \\
 & & \quad = V_2 \Leftrightarrow \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = (1 - \omega_1\omega_2 + \omega_1\omega_2 - \omega_2)/ \\
 & & \quad (1 - \omega_1\omega_2)\&\omega_1 = V_2 \\
 & \omega_1 = V_2 = 0 & \text{OR } V_1 = 1 - (1 - \omega_1) \\
 & & \quad \times\omega_2/(1 - \omega_1\omega_2)\&\omega_1 \\
 & & \quad = V_2
\end{array}
$$

Case $V_1 = 1$ and $\omega_2 = 1$ for the memory traces approach
For $\omega_2 = 1$ this can be simplified as

$$\omega_1 = V_2 = (\beta V_1 - 1 + \beta)/2\beta - 1)V_1 \quad \text{OR} \quad V_2 = 0 \quad \text{OR} \quad \beta = 0.5\ \&\ V_1 = 1$$

and when also $V_1 = 1$ it becomes:

$$\omega_1 = V_2 = (\beta - 1 + \beta)/(2\beta - 1) = 1 \text{ OR } V_2 = 0 \text{ OR } \beta = 0.5\ \&\ V_1 = 1$$

6 Discussion

In this paper a number of learning models for the induction strength of an emotional response on a stimulus were analysed and compared. The introduced models on the one hand describe more specifically how a stimulus generates an emotional response that is felt, and on the other hand how the induction strength of the experienced emotional response is adapted over time. For feeling the emotion, a converging recursive body loop was used, based on elements taken from [1, 2, 6]. One of the

adaptation models was based on a Hebbian learning rule cf. [7–9, 16]. Another one was based on temporal discounting, and the third one was based on memory traces. The models were specified in the hybrid dynamic modelling language LEADSTO, and simulations were performed in its software environment; cf. [14]. Moreover, a mathematical analysis was made to determine possible equilibria. In the comparison differences in adaptation speed and pattern have been found, and in the maximal value of the induction strength.

References

1 Bosse, T., Jonker, C.M., Treur, J.: Formalisation of Damasio′s theory of emotion, feeling and core consciousness. Conscious. Cogn. J., **17** (2008) 94–113
2 Damasio, A.: The Feeling of What Happens. Body and Emotion in the Making of Consciousness. New York, NY: Harcourt Brace (1999)
3 Winkielman, P., Niedenthal, P.M., Oberman, L.M.: Embodied perspective on emotion-cognition interactions. In Pineda, J.A., ed.: Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition. New York, NY: Humana Press/Springer Science (2009) 235–257
4 Damasio, A.: Descartes' Error: Emotion, Reason and the Human Brain. London: Papermac (1994)
5 Keysers, C., Perrett, D.I.: Demystifying social cognition: a Hebbian perspective. Trends Cogn. Sci. **8** (2004) 501–507
6 Damasio, A.: Looking for Spinoza. London: Vintage books (2004)
7 Bi, G.Q., Poo, M.M.: Synaptic modifications by correlated activity: Hebb's postulate revisited. Ann. Rev. Neurosci. **24** (2001) 139–166
8 Gerstner, W., Kistler, W.M.: Mathematical formulations of Hebbian learning. Biol. Cybern. **87** (2002) 404–415
9 Hebb, D.O.: The Organisation of Behavior. New York, NY: Wiley (1949)
10 Bosse, T., Schut, M.C., Treur, J., Wendt, D.: Trust-Based Inter-Temporal decision making: emergence of altruism in a simulated society. In: Antunes, L., Paolucci, M., Norling, E., eds.: Proceedings of the Eighth International Workshop on Multi-Agent-Based Simulation, MABS′07, 2007. Lecture Notes in Artificial Intelligence, vol. 5003. Berlin: Springer Verlag (2008) 96–111
11 Jonker, C.M., Treur, J.: Formal analysis of models for the dynamics of trust based on experiences. In Garijo, F.J., Boman, M., eds.: Multi-Agent System Engineering, Proceedings of the 9th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW′99. Lecture Notes in AI, vol. 1647. Berlin: Springer Verlag (1999) 221–232
12 Shors, T.J.: Memory traces of trace memories: neurogenesis, synaptogenesis and awareness. Trends Neurosci. **27** (2004) 250–256
13 Weinberger, N.M.: Specific long-term memory traces in primary auditory cortex. Nat. Rev. Neurosci. **5** (2004) 279–290
14 Bosse, T., Jonker, C.M., Meij, L. van der, Treur, J.: A language and environment for analysis of dynamics by simulation. Int. J. Artif. Intell. Tools **16** (2007a) 435–464
15 Bosse, T., Jonker, C.M., Treur, J.: Simulation and analysis of adaptive agents: an integrative modelling approach. Adv. Complex Syst. J. **10** (2007b) 335–357
16 Xie, X., Seung, H.S.: Equivalence of Backpropagation and Contrastive Hebbian Learning in a Layered Network. Neural Comput. **15** (2003) 441–454

Modelling Trust Dynamics from a Neurological Perspective

Mark Hoogendoorn, S. Waqar Jaffry, and Jan Treur

Abstract Trust dynamics is often modelled in relation to experiences and usually from a cognitive perspective, leaving affective aspects out of consideration. Inspired by neurological findings that show more and more how in mental processes cognitive and affective aspects are intertwined, in this paper, a model for trust dynamics is introduced incorporating the relation between trust and feeling. The model makes use of a Hebbian learning principle and describes how trust depends on experiences viewed as information obtained over time, but also as emotional responses and feelings.

Keywords Trust dynamics · Hebbian learning · Cognitive and affective

1 Introduction

For a human performing certain task, an important aspect is the trust in various information sources that can potentially aid in performing the task. For instance, the person may have experienced in the past that the manual for the task contains a lot of flaws, whereas it was experienced that a knowledgeable colleague can immediately aid in an adequate fashion. As a result the trust level for the manual will be much lower than the trust level for the colleague, and a support system offering the manual will most likely be ignored. This illustrates two main functional properties of trust states as cognitive states that are often considered [1]:

(1) A trust state accumulates experiences over time
(2) Trust states are used in decision making by choosing more trusted options above less trusted options

M. Hoogendoorn (✉)
Department of Artificial Intelligence, VU University Amsterdam, Amsterdam, 1081 HV, The Netherlands
e-mail: mhoogen@few.vu.nl

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_81,

In trust research, a variety of computational models have been proposed for human trust, see e.g. [1–4]. Such models often assume that trust values over time depend on the experiences of the human with the specific object of trust. In [4] an additional factor is taken into account, namely how the object of trust performs relative to its competitors. Most of such models consider experiences and trust as cognitive concepts and exclude affective factors. This contrasts with how persons in general experience trust and trust-affecting experiences, which, for example, may go hand in hand with strong feelings of disappointment or insecurity. Much work reports interactions between cognitive and affective aspects for a variety of cases [5–10], without relating this explicitly to neurological findings or theories. In the current paper, neurological theories on emotion and feeling are adopted as a basis of inspiration for a computational model for trust dynamics that models such interactions. The computational model, which is inspired by neurological theories on the embodiement of emotions as described, for example, in [10–14], describes how trust dynamics relates to experiences with (external) sources, both from a cognitive and affective perspective. More specifically, in accordance with, for example [13] and [14], for feeling the emotion associated to a mental state, a converging recursive body loop is assumed. In addition, based on Hebbian learning (cf. [15–17]) for the strength of the connections to the emotional responses an adaptation process is introduced, inspired by the Somatic Marker Hypothesis [11, 12]. Compared to more detailed neural models, the model presented here abstracts from more fine-grained descriptions of neurons, and, for example, their biochemistry and spiking patterns. States in the presented model may be viewed as abstract descriptions of states of neurons or rather of groups of neurons. Such a more abstract representation provides a less complex and more manageable model, while the assumption is that it can still show the essential dynamics from a higher level perspective.

In this paper, first in Section 2 Damasio's theory on the generation of feelings is briefly introduced. In Sections 3 and 4 the model is described in detail, including the adaptation based on Hebbian learning. Section 5 presents some simulation results. In Section 6 it is discussed how functional properties of trust states as cognitive states were formulated and automatically verified. Finally, Section 7 is a discussion.

2 On the Interaction Between Affective and Cognitive Aspects

Cognitive states of a person, such as sensory or other representations often induce emotions felt within this person, as described by neurologist Damasio [13] and [14]; for example: "Through either innate design or by learning, we react to most, perhaps all, objects with emotions, however weak, and subsequent feelings, however feeble." [14, p. 93]. In some more detail, emotion generation via a body loop roughly proceeds according to the following causal chain; see [13, 14]:

cognitive state → preparation for bodily response → bodily response →
sensing the bodily response → sensory representation of the bodily response → feeling

The body loop (or as if body loop) is extended to a recursive body loop (or recursive as if body loop) by assuming that the preparation of the bodily response

is also affected by the state of feeling the emotion as an additional causal relation: feeling → preparation for the bodily response. Such recursiveness is also assumed by Damasio [14], as he notices that what is felt by sensing is actually a body state which is an internal object, under control of the person:

> The brain has a direct means to respond to the object as feelings unfold because the object at the origin is inside the body, rather than external to it. (. . .) The object at the origin on the one hand, and the brain map of that object on the other, can influence each other in a sort of reverberative process that is not to be found, for example, in the perception of an external object. [14, p. 91]

Within the model presented in this paper, both the bodily response and the feeling are assigned a level or gradation, expressed by a number. The causal cycle is modelled as a positive feedback loop, triggered by a mental state and converging to a certain level of feeling and body state.

Another neurological theory addressing the interaction between cognitive and affective aspects can be found in Damasio's Somatic Marker Hypothesis; cf. [11, 12, 14, 18]. This is a theory on decision making which provides a central role to emotions felt. Within a given context, each represented decision option induces (via an emotional response) a feeling which is used to mark the option. For example, a strongly negative somatic marker linked to a particular option occurs as a strongly negative feeling for that option. Similarly, a positive somatic marker occurs as a positive feeling for that option. Damasio describes the use of somatic markers in the following way:

> the somatic marker (. . .) forces attention on the negative outcome to which a given action may lead, and functions as an automated alarm signal which says: beware of danger ahead if you choose the option which leads to this outcome. The signal may lead you to reject, *immediately*, the negative course of action and thus make you choose among other alternatives. (. . .) When a positive somatic marker is juxtaposed instead, it becomes a beacon of incentive. . . ([11], pp. 173–174)

Somatic markers may be innate, but may also by adaptive, related to experiences:

> Somatic markers are thus acquired through experience, under the control of an internal preference system and under the influence of an external set of circumstances . . . ([11] p. 179)

In the model introduced below, this adaptive aspect will be modelled as Hebbian learning; cf. [15–17]. Viewed informally, in the first place it results in a dynamical connection strength obtained as an accumulation of experiences over time (1). Secondly, in decision making this connection plays a crucial role as it determines the emotion felt for this option, which is used as a main decision criterion (2). As discussed in the introduction, these two properties (1) and (2) are considered two main functional, cognitive properties of a trust state. Therefore they give support to the assumption that the strength of this connection can be interpreted as a representation of the trust in the option considered.

3 Incorporating Affective Aspects in a Trust Model

Informally described theories in scientific disciplines, for example, in biological or neurological contexts, often are formulated in terms of causal relationships or in terms of dynamical systems. To adequately formalise such a theory the hybrid dynamic modelling language LEADSTO has been developed that subsumes qualitative and quantitative causal relationships, and dynamical systems; cf. [19]. This language has been proven successful in a number of contexts, varying from biochemical processes that make up the dynamics of cell behaviour [20] to neurological and cognitive processes [21–23]. Within LEADSTO a temporal relation a → b denotes that when a state property a occurs, then after a certain time delay (which for each relation instance can be specified as any positive real number), state property b will occur. In LEADSTO both logical and numerical calculations can be specified in an integrated manner; a dedicated software environment is available to support specification and simulation.

An overview of the model for how trust dynamics emerges from the experiences is depicted in Fig. 1. How decisions are made, given these trust states is depicted in Fig. 2. These pictures also show representations from the detailed specifications explained below. However, note that the precise numerical relations between the indicated variables V shown are not expressed in this picture, but in the detailed specifications of properties below, which are labeled by LP1 to LP11 as also shown in the pictures. The detailed specification (both informally and formally) of the model is presented below. Here capitals are used for (assumed universally quantified) variables. First the part is presented that describes the basic mechanisms to generate a belief state and the associated feeling. The first dynamic property addresses how properties of the world state can be sensed.

LP1 Sensing a world state

If world state property W occurs of strength *V*

then a sensor state for W of strength *V* will occur.

```
world_state(W, V) → sensor_state(W, V)
```

Note that this generic dynamic property is used for a specific world state, for experiences with the different options and for body states; to this end the variable W is instantiated respectively by w, exp1 and exp2, b1 and b2. From the sensor states, sensory representations are generated according to the dynamic property LP2. Note that also here for the example the variable P is instantiated as indicated.

LP2 Generating a sensory representation for a sensed world or body state

If a sensor state for world state or body state property P with level *V* occurs,

then a sensory representation for P with level *V* will occur.

```
sensor_state(P, V) → srs(P, V)
```

For a given world state representations for a number of options are activated:

LP3 Generating an option for a sensory representation of a world state

If a sensory representation for w with level *V* occurs

then a representation for optinon o with level *V* will occur

```
srs(w, V) → rep(o, V)
```

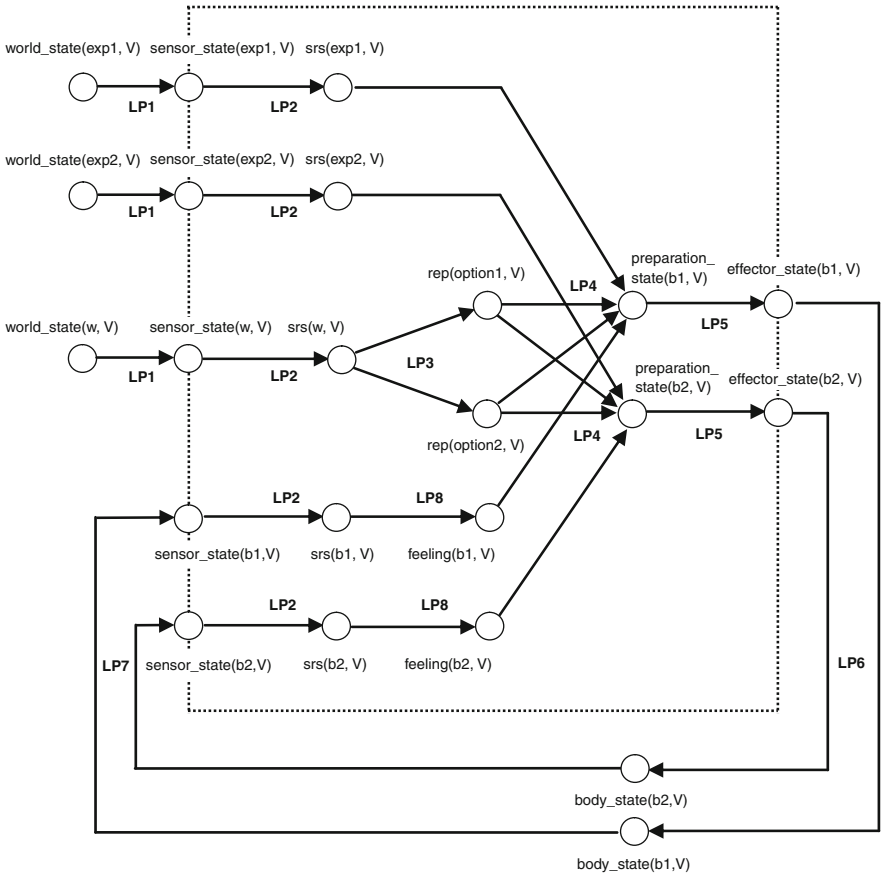


Fig. 1 Overview of the model for trust dynamics

Dynamic property LP4 describes the emotional response to the person's mental state in the form of the preparation for a specific bodily reaction. Here the mental state comprises a number of cognitive and affective aspects: options activated, experienced results of options and feelings. This specifies part of the loop between feeling and body state. This property uses a combination model based on a function $g(\beta_1,\ \beta_2,\ V_1,\ V_2,\ \omega_1,\ \omega_2,\ \omega_3)$ including a threshold function. For example,

$$g(\beta_1,\ \beta_2,\ V_1,\ V_2,\ V_3,\ \omega_1,\ \omega_2,\ \omega_3) = th(\beta_1,\ \beta_2,\ \omega_1 V_1 + \omega_2 V_2 + \omega_3 V_3)$$

with $V_1,\ V_2,\ V_3$ activation levels and $\omega_1,\ \omega_2,\ \omega_3$ weights of the connections, and threshold function $th(\beta_1,\ \beta_2, V) = 1/(1 + e^{-\beta_2(V-\beta_1)})$ with threshold β_1 and steepness β_2.

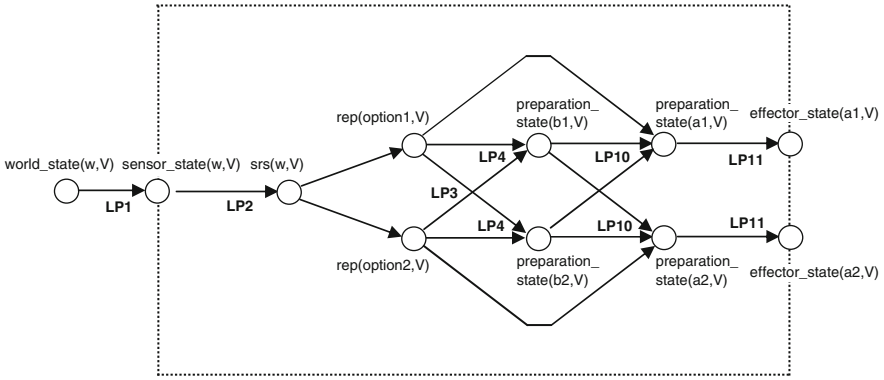


Fig. 2 Overview of the model for trust-based decision making

LP4a From option activation and experience to preparation of a body state (non-competitive case)

If option o with level V_1 occurs
and feeling the associated body state b has level V_2
and an experience for o occurs with level V_3
and the preparation state for b has level V_4
then a preparation state for body state b will occur with level $V_4 + \gamma(g(\beta_1, \beta_2, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3) - V_4)\,\Delta t$

$$\texttt{rep(o, V}_1\texttt{) \& feeling(b, V}_2\texttt{) \& srs(exp, V}_3\texttt{) \& preparation\_state(b, V}_4\texttt{)} \rightarrow \texttt{preparation\_state(b, V}_4 + \gamma(g(\beta_1, \beta_2, V_1, V_2, V_3, \omega_1, \omega_2, \omega_3) - \texttt{V}_4)\Delta\texttt{t)}$$

For the competitive case also the inhibiting cross connections from one represented option to the body state induced by another represented option are used. A function involving these cross connections was defined, for example

$$h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) = th(\beta_1, \beta_2, \omega_1 V_1 + \omega_2 V_2 + \omega_3 V_3 - \omega_{21} V_{21})$$

for two considered options, with ω_{21} the weight of the suppressing connection from represented option 2 to the preparation state induced by option 1.

LP4b From option activation and experience to preparation of a body state (competitive case)

If option o1 with level V_1 occurs
and option o2 with level V_{21} occurs
and feeling the associated body state b1 has level V_2
and an experience for o1 occurs with level V_3
and the preparation state for b1 has level V_4
then a preparation state for body state b1 will occur with
level $V_4 + \gamma(h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21}, \omega_1, \omega_2, \omega_3, \omega_{21}) - V_4)\Delta t$.

$$\texttt{rep(o1, V}_1\texttt{) \& rep(o2, V}_{21}\texttt{) \& feeling(b1, V}_2\texttt{) \& srs(exp1, V}_3\texttt{) \& preparation\_state(b1, V}_4\texttt{)} \rightarrow \texttt{preparation\_state(b1, V}_4 + y(h(\beta_1, \beta_2, V_1, V_2, V_3, V_{21,\omega_1, \omega_2, \omega_3, \omega_{21}}) - V_4)\Delta\texttt{t)}$$

Dynamic properties LP5, LP6, and LP7 together with LP2 describe the body loop.

LP5 From preparation to effector state for body modification
If preparation state for body state B occurs with level V,
then the effector state for body state B with level V will occur.
`preparation_state(B, V) → effector_state(B, V)`
LP6 From effector state to modified body state
If the effector state for body state B with level V occurs,
then the body state B with level V will occur.
`effector_state(B, V) → body_state(B, V)`
LP7 Sensing a body state
If body state B with level V occurs,
then this body state B with level V will be sensed.
`body_state(B, V) → sensor_state(B, V)`
LP8 From sensory representation of body state to feeling
If a sensory representation for body state B with level V occurs,
then B is felt with level V.
`srs(B, V) → feeling(B, V)`

Alternatively, dynamic properties LP5 to LP7 can also be replaced by one dynamic property LP9 describing an as if body loop as follows.

LP9 From preparation to sensed body state
If preparation state for body state B occurs with level V,
then the effector state for body state B with level V will occur.
`preparation_state(B, V) → srs(B, V)`

For the decision process on which option O_i to choose, represented by action A_i, a winner-takes-it-all model is used based on the feeling levels associated to the options; for an overview, see Fig. 2. This has been realised by combining the option representations O_i with their related emotional responses B_i in such a way that for each i the level of the emotional response B_i has a strongly positive effect on preparation of the action A_i related to option O_i itself, but a strongly suppressing effect on the preparations for actions A_j related to the other options O_j for $j \neq i$. As before, this is described by a function

$$h(\beta_1,\ \beta_2,\ V_1,\ \ldots,\ V_m,\ U_1,\ \ldots,\ U_m,\ \omega_{11},\ \ldots,\ \omega_{mm})$$

with V_i levels for representations of options O_i and U_i levels of preparation states for body state B_i related to options O_i and ω_{ij} the strength of the connection between preparation states for body state B_i and preparation states for action A_j.

LP10 Decisions based on felt emotions induced by the options
If options O_i with levels V_i occur,
and preparation states for body state B_i related to options O_i occur with level U_i,
and the preparation state for action A_i for option O_i has level W_i
then the preparation state for action A_i for option O_i will occur
with level $W_i + \gamma(h(\beta_1,\ \beta_2,\ V_1, \ldots,\ V_m,\ U_1,\ \ldots,\ U_m, \omega_{11},\ ..\ \omega_m) - W_i)\ \Delta t$
`rep(O1, V1) & ... & rep(Om, Vm) &`
`preparation_state(B1, U1) & ... & preparation_state(Bm, Um) &`
`preparation_state(Ai, Wi)`
→ `preparation_state(Ai, Wi +` $\gamma(h(\beta_1,\ \beta_2,\ V_1,\ \ldots,\ V_m,\ U_1,\ \ldots,\ U_m, \omega_{11},\ ..\ \omega_{mm}) - W_i)\ \Delta t)$

LP11 From preparation to effector state for an action
If preparation state for action A occurs with level V,
then the effector state for action A with level V will occur.
`preparation_state(A, V)` → `effector_state(A, V)`

4 The Hebbian Adaptation Process

From a neurological perspective the strength of a connection from an option to an emotional response may depend on how experiences are felt emotionally, as neurons involved in the option, the preparation for the body state, and in the associated feeling will often be activated simultaneously. Therefore such a connection from option to emotional response may be strengthened based on a general Hebbian learning mechanism [15–17] that states that connections between neurons that are activated simultaneously are strengthened, similar to what has been proposed for the emergence of mirror neurons; e.g. [24, 25]. This principle is applied to the strength ω_1 of the connection from option 1 to the emotional response expressed by body state b1. The following learning rule takes into account a maximal connection strength 1, a learning rate η, and an extinction rate ζ.

LP12 Hebbian learning rule for the connection from option to preparation
If the connection from option o1 to preparation of b1 has strength ω_1
and the option o1 has strength V_1
and the preparation of b1 has strength V_2
and the learning rate from option o1 to preparation of b1 is η
and the extinction rate from option o1 to preparation of b1 is ζ
then after Δt the connection from option o1 to preparation state b1 will have strength $\omega_1 + (\eta V_1 V_2(1 - \omega_1) - \zeta\omega_1)\,\Delta t$.
`has_connection_strength(rep(o1), preparation(b1), ω1) & rep(o1, V1) & preparation(b1, V2) & has_learning_rate(rep(o1), preparation(b1), η) & has_extinction_rate(rep(o1), preparation(b1), ζ)` → `has_connection_strength(rep(o1), preparation(b1), (b1), ω1+(ηV1V2(1−ω1)−ζω1)Δt)`

By this rule through their affective aspects, the experiences are accumulated in the connection strength from option o1 to preparation of body state b1, and thus serves as a representation of trust in this option o1. A similar Hebbian learning rule can be found in ([17] p. 406).

5 Example Simulation Results

The model described in Section 3 has been used to generate a number of simulation experiments for non-competitive and competitive cases (see Fig. 3 for some example results). To ease the comparison between these cases the same model parameter values were used for these examples (see Table 1). In Fig. 3a example simulation results are shown for the non-competitive case. Here the subject is exposed to an information source that provides experience values 0 respectively 1 alternating periodically in a period of 200 time steps each. In this figure it can be observed that change in

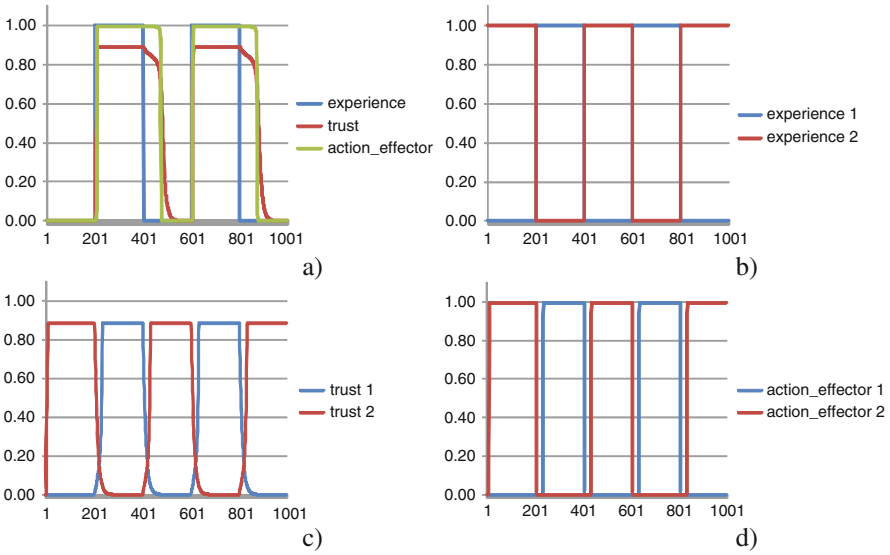


Fig. 3 Simulation results: experience, connection representing trust, and action effector state for a non-competitive case in (**a**) and a competitive case in (**b**), (**c**) and (**d**) respectively

Table 1 Parameter values used in the example simulations

Parameter	Value	Meaning
β_1	0.95	Threshold value for preparation state and action effector state
β_2	10, 100	Steepness value for preparation state, action effector state
γ	0.90	Activation change rate
η	0.80	Learning rate of connection from option representation to preparation
ζ	0.10	Extinction rate of connection from option representation to preparation
Δt	0.90	Time step
ω_s, ω_a	0.50	Suppressing weight from option representation to preparation state and from preparation state to the action state (competitive case)

experience leads to changes in the connection strengths (representing trust) as well as the action effector states. Furthermore, the decrease in the connection strengths representing trust due to a bad experience (0) takes longer than the increase due to a good experience (1), which can be explained by the higher value of the learning rate than of the extinction rate.

In Figs. 3b–d, the simulation results are shown for the competitive case with two competitive options having suppression weight 0.5 from option representation to preparation state and from preparation state to the action state. In this case the subject is exposed to two information sources that provides experience values 0

respectively 1 alternating periodically in a period of 200 time steps each, in a reverse cycle with respect to each other (see Fig. 3b). Here change in experience changes the connections representing trust as well as the action effector states. Moreover, in comparison to the non-competitive case, the learning is slow while decay is fast, which is due to the presence of competition. Finally Fig. 3 shows that the connection strengths in the presented model exhibit the two fundamental functional properties of trust discussed in Section 1, namely that trust is based on accumulation of experiences over time (see Fig. 3c) and that trust states are used in decision making by choosing more trusted options above less trusted ones (see Fig. 3d).

6 Verification of Functional Properties of Trust

The two functional properties of trust states formulated in the introduction are: (1) A trust state accumulates experiences over time, and (2) Trust states are used in decision making by choosing more trusted above less trusted options. These properties characterise trust states from a functional, cognitive perspective. Therefore any model or computational or physical realisation claimed to describe trust dynamics has to (at least) satisfy these properties. Such properties can be formalized in a temporal logical language, and can be automatically verified for the traces that have been generated using the proposed model. In this section this verification of properties is discussed. First, the language used to verify these properties is explained. Thereafter the properties and the results of the verification are discussed.

The verification of properties has been performed using a language called TTL (Temporal Trace Language), that features a dedicated editor and an automated checker; cf. [26]. This predicate logical temporal language supports formal specification and analysis of dynamic properties, covering both qualitative and quantitative aspects. TTL is built on atoms referring to *states* of the world, *time points* and *traces*, i.e. trajectories of states over time. In addition, *dynamic properties* are temporal statements that can be formulated with respect to traces based on the state ontology Ont in the following manner. Given a trace γ over state ontology Ont, the state in γ at time point t is denoted by state(γ, t). These states can be related to state properties via the infix predicate |=, where state(γ, t) |= p denotes that state property p holds in trace γ at time t. Based on these statements, dynamic properties can be formulated using quantifiers over time and traces and the usual first-order logical connectives such as $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\forall$, $\exists$. For more details, see [26].

In order to be able to automatically verify the properties upon the simulation traces, they have been formalised. From the computational verification process it was found that indeed they are satisfied by the simulation traces of the model for which they were verified. The first functional property (1), specifying that a trust state accumulates experiences over time, is split up into a number of properties. First, two properties are specified which express trust accumulation for the non-competitive case, whereby the connections for the respective trustees are not influenced by experiences with competitors.

P1.1 Connection strength increases with more positive experience (non-competitive case)
If a sensor state indicates a particular value V_1 of an experience E, and E is an experience for trustee T, and the current strength of the connection for trustee T is V_2, and V_1 is higher than V_2, then the connection strength will remain the same or increase.

```
∀γ:TRACE, t:TIME, E:EXPERIENCE, T:TRUSTEE, V1,V2,V3:VALUE
[ state(γ, t) |= sensor_state(E, V1) & state(γ, t) |= connection(T, V2) &
 state(γ, t) |= matches(E, T) & V1 > V2
 ⇒ ∃V3:VALUE [ state(γ, t+1) |= connection(T, V3) & V3 ≥ V2 ] ]
```

P1.2 Connection strength decreases with more negative experience (non-competitive case)
If a sensor state indicates a particular value V_1 of an experience E, and E is an experience for trustee T, and the current strength of the connection for trustee T is V_2, and V_1 is lower than V_2, then the connection strength will remain the same or decrease.

```
∀γ:TRACE, t:TIME, E:EXPERIENCE, T:TRUSTEE, V1,V2,V3:VALUE
[ state(γ, t) |= sensor_state(E, V1) & state(γ, t) |= connection(T, V2) &
 state(γ, t) |= matches(E, T) & V1 < V2
 ⇒ ∃V3:VALUE [ state(γ, t+1) |= connection(T, V3) & V3 ≤ V2 ] ]
```

Besides the non-competitive case, also properties have been specified for the competitive case. Hereby, the experiences with other competitive information sources are also taken into account.

P2.1 Connection strength increases with more positive experience (competitive case)
If a sensor state indicates a particular value V_1 of an experience E, and E is an experience for trustee T, and the current strength of the connection for trustee T is V_2, and V_1 is higher than V_2, and all other experiences are lower compared to V_1, then the connection strength will remain the same or increase.

```
∀γ:TRACE, t:TIME, E:EXPERIENCE, T:TRUSTEE, V1,V2,V3:VALUE
[ state(γ, t) |= sensor_state(E, V1) &
∀E' ≠ E [ ∃V':VALUE state(γ, t) |= sensor_state(E', V') & V' < V1 ] &
 state(γ, t) |= connection(T, V2) & state(γ, t) |= matches(E, T) & V1 > V2
 ⇒ ∃V3:VALUE [ state(γ, t+1) |= connection(T, V3) & V3 ≥ V2 ] ]
```

P2.2 Connection strength decreases with more negative experience (competitive case)
If a sensor state indicates a particular value V_1 of an experience E, and E is an experience for trustee T, and the current strength of the connection for trustee T is V_2, and V_1 is lower than V_2, and all other experiences are higher compared to V_1, then the connection strength will remain the same or decrease.

```
∀γ:TRACE, t:TIME, E:EXPERIENCE, T:TRUSTEE, V1,V2,V3:VALUE
[ state(γ, t) |= sensor_state(E, V1) &
∀E' ≠ E [ ∃V':VALUE state(γ, t) |= sensor_state(E', V') & V' >V1 ] &
 state(γ, t) |= connection(T, V2) & state(γ, t) |= matches(E, T) & V1 < V2
 ⇒ ∃V3:VALUE [ state(γ, t+1) |= connection(T, V3) & V3 ≤ V2 ] ]
```

Finally, property P3 is specified which compares different traces, as shown below.

P3.1 Higher experiences lead to higher connection strengths (non-competitive case)
If within one trace the experiences for a trustee are always higher compared to the experiences for a trustee in another trace, then in that trace the connection strengths will always be higher.

```
∀γ1, γ2:TRACE, E:EXPERIENCE, T:TRUSTEE
[γ1≠ γ2 & state(γ1, 0) |= matches(E, T) &
∀t:TIME [ ∃V1, V2:VALUE [ state(γ1, t) |= sensor_state(E, V1) &
 state(γ2, t) |= sensor_state(E, V2) & V1>V2 ] ]
```

```
⇒ ∀t:TIME [ ∃V1, V2:VALUE [ state(γ1, t) |= connection(T, V1) &
              state(γ2, t) |= connection(T, V2) & V1>V2 ] ] ]
```

P3.2 Higher experiences lead to higher connection strengths (competitive case)
If within one trace the experiences for a trustee are always higher compared to the experiences for a trustee in another trace, and there are no other experiences with a higher value at that time point, then in that trace the connection strengths will always be higher.

```
∀γ1, γ2:TRACE, E:EXPERIENCE, T:TRUSTEE
[γ1≠ γ2 & state(γ1, 0) |= matches(E, T) &
   ∀t:TIME [ ∃V1, V2:VALUE [ state(γ1, t) |= sensor_state(E, V1) &
    ∀E' ≠ E [ ∃V':VALUE state(γ1, t) |= sensor_state(E', V') & V' ≤ V1 ] &
    state(γ2, t) |= sensor_state(E, V2) &
    ∀E'' ≠ E [ ∃V'':VALUE state(γ2, t) |= sensor_state(E'', V'') & V'' ≤ V2 ]
   & V1>V2 ] ]
⇒ ∀t:TIME [ ∃V1, V2:VALUE [ state(γ1, t) |= connection(T, V1) &
    state(γ2, t) |= connection(T, V2) & V1 > V2 ] ] ]
```

The formalization of the second functional property (2), i.e., trust states are used in decision making by choosing more trusted options above less trusted options, is expressed as follows.

P4 The trustee with the strongest connection is selected
If for trustee T the connection strength is the highest, then this trustee will be selected.

```
∀γ:TRACE, t1:TIME, T:TRUSTEE, V1:VALUE
[[ state(γ, t1) |= connection(T, V1) & state(γ, t1) |= sensor_state(w, 1) &
  ∀T', V' [ [ T' ≠ T & state(γ, t1) |= connection(T', V') ] ⇒ V' < V1 ] ]
  ⇒ ∃t2:TIME < t1 + 10  [ state(γ, t2) |= effector_state(T) ]
```

Note that in the property, the effector state merely has one argument, namely the trustee with the highest effector state value.

7 Discussion

In this paper a computational model for trust dynamics was introduced incorporating the reciprocal interaction between cognitive and affective aspects based on neurological theories that address the role of emotions and feelings. The introduced model describes more specifically how considered options and experiences generate an emotional response that is felt. For feeling the emotion, based on elements taken from [13, 14, 23], a converging recursive body loop is included in the model. An adaptation process based on Hebbian learning [15–17], was incorporated, inspired by the Somatic Marker Hypothesis described in [11, 12, 18], and as also has been proposed for the functioning of mirror neurons; e.g. [24, 25]. The model was specified in the hybrid dynamic modelling language LEADSTO, and simulations were performed in its software environment; cf. [19].

It has been shown that within the model the strength of the connection between a considered option and the emotional response induced by it, satisfies two properties that are considered as two central functional properties of a trust state as a cognitive state [1]: (1) it accumulates experiences, and (2) it is a crucial factor used in deciding for the option. This provides support for the assumption that the strength of this

connection can be interpreted as a representation of the trust level in the considered option.

Models of neural processes can be specified at different levels of abstraction. The model presented here can be viewed as a model at a higher abstraction level, compared to more detailed models that take into account more fine-grained descriptions of neurons and their biochemical and/or spiking patterns. States in the presented model can be viewed as abstract descriptions of states of neurons or as representing states of groups of neurons. An advantage of a more abstract representation is that such a model is less complex and therefore may be less difficult to handle, while it still shows the essential dynamics. An interesting issue for further research is how such a more abstract model can be related to more detailed models, and in how far patterns observed in more specific models also are represented in such a more abstract model.

References

1. Jonker, C.M., Treur, J.: A temporal-interactivist perspective on the dynamics of mental states. Cogn. Syst. Res. J. **4** (2003) 137–155
2. Jonker, C.M., Treur, J.: Formal Analysis of Models for the Dynamics of Trust based on Experiences, In Garijo, F.J., Boman, M., eds.: Proceedings of MAAMAW′99, LNAI, vol. 1647. Berlin: Springer Verlag (1999) 221–232
3. Falcone, R., Castelfranchi, C.: Trust dynamics: how trust is influenced by direct experiences and by trust itself. In Proceedings of AAMAS 2004. (2004) 740–747
4. Hoogendoorn, M., Jaffry, S.W., Treur, J.: Modeling Dynamics of relative trust of competitive information agents. In Klusch, M., Pechoucek, M., Polleres, A., eds.: Proceedings of the 12th International Workshop on Cooperative Information Agents, CIA′08. LNAI, vol. 5180. Berlin: Springer (2008) 55–70
5. Eich, E., Kihlstrom, J.F., Bower, G.H., Forgas, J.P., Niedenthal, P.M.: Cognition and Emotion. New York, NY: Oxford University Press (2000)
6. Forgas, J.P., Laham, S.M., Vargas, P.T. (2005). Mood effects on eyewitness memory: affective influences on susceptibility to misinformation. J. Exp. Social Psychol. **41** (2005) 574–588
7. Forgas, J.P., Goldenberg, L., Unkelbach, C.: Can bad weather improve your memory? An unobtrusive field study of natural mood effects on real-life memory. J. Exp. Social Psychol. **45** (2009) 254–257
8. Niedenthal, P.M.: Embodying Emotion. Science **316** (2007) 1002–1005
9. Schooler, J.W., Eich, E.: Memory for Emotional Events. In Tulving, E., Craik, F.I.M., eds.: The Oxford Handbook of Memory. Oxford: Oxford University Press (2000) 379–394
10. Winkielman, P., Niedenthal, P.M., Oberman, L.M.: Embodied perspective on emotion-cognition interactions. In Pineda, J.A., ed.: Mirror Neuron Systems: the Role of Mirroring Processes in Social Cognition. New York, NY: Humana Press/Springer Science (2009) 235–257
11. Damasio, A.: Descartes' Error: Emotion, Reason and the Human Brain. London: Papermac (1994)
12. Damasio, A.: The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex. Philos. Trans. R. Soc. Biol. Sci. **351** (1996) 1413–1420
13. Damasio, A.: The Feeling of What Happens. Body and Emotion in the Making of Consciousness. New York, NY: Harcourt Brace (1999)
14. Damasio, A.: Looking for Spinoza. London: Vintage books (2004)
15. Hebb, D.: The Organisation of Behavior. New York, NY: Wiley (1949)

16. Bi, G.Q., Poo, M.M.: Synaptic modifications by correlated activity: Hebb's postulate revisited. Ann. Rev. Neurosci. **24** (2001) 139–166
17. Gerstner, W., Kistler, W.M.: Mathematical formulations of Hebbian learning. Biol. Cybern. **87** (2002) 404–415
18. Bechara, A., Damasio, A.: The Somatic marker hypothesis: a neural theory of economic decision. Games Econ. Behav. **52** (2004) 336–372
19. Bosse, T., Jonker, C.M., Meij, L. van der, Treur, J.: A language and environment for analysis of dynamics by simulation. Int. J. Artif. Intell. Tools **16** (2007) 435–464
20. Jonker, C.M., Snoep, J.L., Treur, J., Westerhoff, H.V., Wijngaards, W.C.A.: BDI-modelling of complex intracellular dynamics. J. Theor. Biol. **251** (2008) 1–23
21. Bosse, T., Jonker, C.M., Los, S.A., Torre, L. van der, Treur, J.: Formal analysis of trace conditioning. Cogn. Syst. Res. J. **8** (2007) 36–47
22. Bosse, T., Jonker, C.M., Treur, J.: Simulation and analysis of adaptive agents: an integrative modelling approach. Adv. Complex Syst. J. **10** (2007) 335–357
23. Bosse, T., Jonker, C.M., Treur, J.: Formalisation of Damasio′s theory of emotion, feeling and core consciousness. Conscious. Cogn. J. **17** (2008) 94–113
24. Keysers, C., Perrett, D.I.: Demystifying social cognition: a Hebbian perspective. Trends Cogn. Sci. **8** (2004) 501–507
25. Keysers, C., Gazzola, V.: Unifying social cognition. In Pineda, J.A., ed.: Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition. Berlin: Humana Press/Springer Science (2009) 3–28
26. Bosse, T., Jonker, C.M., Meij, L. van der, Sharpanskykh, A., Treur, J.: Specification and verification of dynamics in agent models. Int. J. Cooperative Inf. Syst. **18** (2009) 167–193

Novel Bionic Model for Pattern Recognition

Eduardo González, Yusely Ruiz, and Guang Li

Abstract In this paper, a novel bionic model and it performance to pattern recognition is researched. The model consists of a bulb model and a 3-layered cortical model; mimicking the main features of the olfactory system. The bulb and cortical models are connected by feed forward and feedback fibers with distributed delays. The Iris dataset from UCI repository contains 3 classes of 50 instances each, where each class refers to a type of iris plant was classified. The performance of the novel model was compared with previously published methods.

Keywords Olfactory system · Bionic model · Pattern recognition

1 Introduction

The olfactory system is a relatively simple and well known sensory system that provides an interesting tool for understanding the process of recognition, classification, learning and recall [1, 2]. Mammalian odor information is treated by olfactory circuits developed during millions of years. It could be divided in three main parts: the olfactory epithelium, the olfactory bulb and the olfactory cortex. The odorous molecules are detected by the odor receptors in the epithelium and transduced into neural activation that is sent to the bulb and from there to the olfactory cortex. In this paper, a mathematical bionic model is shown and its capacity to recognize patterns from a multiclass dataset is studied. The Iris dataset from UCI repository, widely used by other researchers, was selected to carry on the experiments. In the next sections a description of the model and the classification results for the Iris dataset are shown.

E. González (✉)
Biomedical Engineering Department, Zhejiang University, Hangzhou 310027, China
e-mail: moreira@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_82,

2 Model Description

On one hand, for comparisons between experiments and theory of the olfactory system, it is always essential to model with realism. On the other hand, to carry out the mathematical analysis and simulation necessary to understand collective and statistical properties, it is required to disregard unnecessary details. Our model organization is a tradeoff between these two considerations. The full model is constituted by a bulbar model which emulate the excitatory mitral and inhibitory granule cells, and of a 3-layered cortical model, which mimics the structure and behavior of the piriform cortex. In our bionic model the feedforward and feedback channel numbers between bulb model and cortical model are equal.

2.1 Network Dynamics

The olfactory bulb can be viewed as the first central olfactory relay station extracting specific stimulus features [3]. In this work, the olfactory bulb was modeled using a simple approximation of excitatory mitral and inhibitory granule cells. The dynamics of each neural ensemble is described using a second order differential equation, based on physiological experiments of the olfactory system [2]:

For mitral and granule layers:

$$\frac{dm_t}{dt} = -am_i + \sum_{j=1}^{N} H_{oij} Q_j \left[g_j(t - \delta_{ij}) \right] + \sum_{j=1}^{N} M_{oij} Q_j \left[m_j(t - \delta_{ij}) \right] + I_{ei}(t) + r \tag{1}$$

$$\frac{dg_i}{dt} = -bg_i + \sum_{j=1}^{N} W_{oij} Q_j \left[m_j(t - \delta_{ij}) \right] + cl_{ci}(t) + r \tag{2}$$

and:

$$\begin{aligned} Q_j(x) &= S'_x \left[\tanh\left(\frac{x - T_h}{S'_x} \right) \right], \textit{if } x \leq T_h \textit{ then } S'_x = 10 S_x \\ Q_j(x) &= S'_x \left[\tanh\left(\frac{x - T_h}{S'_x} \right) \right], \textit{if } x > T_h \textit{ then } S'_x = 7 S_x \end{aligned} \tag{3}$$

where a, b and c represent rate constants. The external input to the mitral layer is symbolized by I_e and I_c represent the cortical feedback input to granule layer. $m_i(t)$ and $g_j(t)$ symbolize the dynamic state of the ith and jth neural ensemble in the mitral and granule layer, respectively. The positive, negative or zero value of the connection strength, M_{oij}, H_{oij}, and W_{oij}, represents connection weights into mitral layer and between mitral and granule layers from neural population j to i; therefore, these matrix values define the system topology. $Q(x)$ is the asymmetric sigmoidal input/output transformation function used for mitral and granule cells, which was derived from the Hodgkin-Huxley equations and S_x represents the maximum asymptote of the sigmoid function [4]. δ correspond to the specific axonal transmission delay and r represent the noise or spontaneous neural activity.

In the cortical model, based on the 3-layered structure [5], all connections were modeled with distance dependent time delays for signal propagation, corresponding to the geometry and fiber characteristics of the real cortex. The dynamics of a network unit is denoted by the mean membrane potential, u, and is given by a set of coupled nonlinear first-order differential delay equations:

$$\frac{du_i}{dt} = -\frac{u_i}{\tau_i} + \sum_{j=1}^{N} w_{ij} g_j[u_j(t - \delta_{ij})] + I_i(t) + r \tag{4}$$

where τ characterizes the spontaneous decay, w_{ij} is the connection weight and δ correspond to the specific axonal transmission delay. The external input to the upper and middle cortical layer from mitral layer is represented by I and r is noise or neural spontaneous activity. The input-output relationship of a neuron populations in the piriform cortex, $g_j(u_j)$, is a continuous sigmoid function, determined by recording evoked potentials, and essentially similar shape to Eq. (3) for the bulb model:

$$g_j = CQ_i \left\{ 1 - \exp\left[\frac{-\exp(u_j) - 1}{Q_i} \right] \right\} \tag{5}$$

where Q_i is the gain parameter and C is a normalization constant.

2.2 Network Learning

For our bionic model, two learning processes are used: Hebbian associative learning and habituation. Hebbian learning under reinforcement set up the memory basins and attractors, while habituation reduces the impact of environment noise. The learning processes are applied in our model to the bulb model and to cortical model.

By the input of the features of each class, the model is trained to discriminate between different classes of patterns. During the training stage the model output, taken from the cortical middle layer energy, determine a cluster for each class and it gravity center. When a test pattern is given, the Euclidian distance from the model output to each gravity center is calculated and the minimum distance determines the cluster to which the pattern belongs.

3 Results

In order to demonstrate the pattern classification performance of the bionic model, the Iris dataset, available from the UCI Machine Learning Data Repository [6], was selected. The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant (Iris Setosa, Iris Versicolour and Iris Virginica). The main reason for choosing this specific dataset is that it is available for public download and it has been studied by other researchers making the results comfortably comparable.

Table 1 Summary of the best classifications for Iris dataset

Methods	Correct classification (%)
Wu and Chen's method [7]	96.210
Castro's method [8]	96.600
Hsu and Lin's method [9]	97.333
Brunzell and Eriksson's method [10]	98.000
Bionic model	98.088

In order to find the optimal number of feedforward channels arriving with signal information from the bulb to the cortex, the number of channels was varied from a minimum of 2.5% of the cortical surface to a maximum of 100% of cortical area. Furthermore, the channels were placed equidistantly into the cortical layers and each configuration was tested fifteen times. The best classification rate was used as indicator of the optimal channel number (88 channels in our computer simulation). The results achieved by our bionic model are compared with previous reported methods which also used the Iris dataset during the validation processes (Table 1).

4 Conclusions

In this work, a novel model mimicking the main features of the olfactory system has been presented. Also the improved performance of the model on the Iris dataset classification, with respect to the results obtained by other researcher's methods during the last decade, was demonstrated. The high classification rate gathered after the computational simulation showed the capacity of our bionic model to learn and classify patterns. Our conclusion is therefore that our bionic model can achieves good results as a pattern classifier on the basis of the structure of mammalian olfactory system.

Acknowledgement The authors would like to tanks professor Hans Liljenström for his helpful comments. This research is supported by the Natural Science Foundation of China (Grants No. 60874098 and 60911130129), the National High Technology Research and Development Program of China (863 Program 2007AA042103) and the National Creative Research Groups Science Foundation of China (NCRGSFC: 60421002).

References

1. Shepherd, G.M.: The Synaptic Organization of the Brain. New York, NY: Oxford University Press (1979).
2. Freeman, W.J.: Mass Action in the Nervous System. New York, NY: Academic Press (1975).
3. Lowe, G.: Electrical signaling in the olfactory bulb. Curr. Opin. Neurobiol. **13**(4) (2003) 476–481.
4. Freeman, W.J.: Nonlinear gain mediating cortical stimulus-response relations. Biol. Cybern **33** (1979) 237–247.

5. Liljenstrom, H.: Neural stability and flexibility: a computational approach. Neuropsychopharmacology **28** (2003) S64–S73.
6. Fisher, R.A.: Iris Data Set. (1936) Available: http://archive.ics.uci.edu/ml/datasets/Iris
7. Wu, T.P., Chen, S.M.: A new method for constructing membership functions and fuzzy rules from training examples. IEEE Trans. Syst. Man Cybern. B Cybern. **29** (1999) 25–40.
8. Castro, J.L., Castro-Schez, J.J., Zurita, J.M.: Learning maximal structure rules in fuzzy logic for knowledge acquisition in expert systems. Fuzzy Sets Syst. **101**(3) (1999) 331–342.
9. Hsu, C.W., Lin, C.J.: A comparison of methods for multiclass support vector machines. IEEE Trans. Neural Netw. **13**(2) (2002) 415–425.
10. Brunzell, H., Eriksson, J.: Feature reduction for classification of multidimensional data. Pattern Recognit. **33**(10) (2000) 1741–1748.

Local Self-Adaptation Mechanisms for Large-Scale Neural System Building

M. Garcia Ortiz and A. Gepperth

Abstract For integrating neural networks into large systems, dynamical stability and parameter settings are key issues, especially for popular recurrent network models such as dynamic neural fields. In neural circuits, homeostatic plasticity seems to counter these problems. Here we present a set of gradient adaptation rules that autonomously regulate the strength of synaptic input and the parameters of the transfer function for each neuron individually. By doing this, we actively maintain the average membrane potentials and firing rates as well as the variances of the firing rate at specified levels. A focus of this contribution lies on clarifying at which time scales these mechanisms should work. The benefit of such self-adaptation is a significant reduction of free parameters as well as the possibility to connect a neural field to almost arbitrary inputs since dynamical stability is actively maintained. We consider these two properties to be crucial since they will facilitate the construction of large neural systems significantly.

Keywords Neural fields · Self-adaptation · Dynamic stability

1 Introduction

It is well known that even single neurons are complex, nonlinear dynamical systems (see, e.g. [1]). Furthermore, neurons are massively interconnected with other neurons by (possibly recurrent) synaptic connections, with their own nonlinear behavior. To maintain dynamical stability under such circumstances, there exist a multitude of activity control mechanisms [2, 1, 3], which autonomously adapt the processing parameters of each neuron according to local rules. These mechanisms are collectively denoted as "homeostatic plasticity". Popular neuron models

M.G. Ortiz (✉)
Honda Research Institute Europe GmbH, 63073 Offenbach, Germany
e-mail: garciaom@minatec.inpg.fr

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_83,

are facing similar stability problems, especially when researchers construct large-scale neural systems [4–7]. In addition, most neuron models contain a multitude of free parameters which cannot always be related to experimental findings.

For both reasons and based on our previous experiences with large-scale neural systems [4], this contribution focuses on local activity control mechanisms for a well-known network model, the dynamic neural field model [8–10]. As a first step, we show how the membrane potentials (related to synaptic scaling [3]) and the firing rates (related to intrinsic plasticity [11, 12]) can be autonomously regulated in face of variable afferent input, thereby maintaining the temporal mean and variance of individual membrane potentials or firing rates at specified target values. Due to the nonlinear dynamic character of the neural field model, it is important to determine time scales such as to minimize interference between adaptation mechanisms.

Several mechanisms of homeostatic plasticity have been previously modeled: on the one hand, intrinsic plasticity, the adaptation of the intrinsic excitability of a neuron's membrane, has been modeled as an adaptation of neural transfer functions in [13, 14] and applied to a number of problems such as reservoir computing [15] or self-organization of sensory representations [16]. On the other hand, authors have modeled synaptic scaling [17], the activity-dependent modification of synaptic strengths [2].

Overall, our work differs from related work in two respects. Firstly, in addition to modeling synaptic scaling and intrinsic plasticity, we demonstrate the operation of these mechanisms concurrently with each other and secondly, we present a strategy of decoupled time scales to prevent interference.

More precisely, the main difference between our work and [13, 16, 14] is the use of recurrent neural networks with a dynamic state instead of input–output encoding neurons. Not needing to address stability problems, these articles strongly focus on achieving a certain output distribution for each neuron. In contrast, we emphasize the reduction in the number of free parameters as well as the dynamic stability. In this respect, our work is related to [17] which also employs the dynamic neural field model, although the focus of our work is on the effects of activity control rather than on self-organization processes.

2 Dynamic Neural Fields

The dynamic neural field model [8] has been proposed to describe pattern formation in the visual cortex. Essentially, dynamic neural fields are a class of recurrent neural network models that have been extensively used for modeling cognitive phenomena like decision making [10], motor planning [9], spatial cognition [18], eye movement preparation [19, 20] and object recognition [21, 22]. Basic elements are simple dynamic-state neurons, a fixed lateral connectivity, and a (usually sigmoid) nonlinearity.

In the neural field model described in [8], natural candidates for self-adaptation are the strengths of afferent inputs and the transfer function parameters, which need to be made position and time dependent for this purpose. We thus formulate a generalized version of the original model suitable for local adaptation mechanisms:

$$\begin{aligned} \tau \dot{u}(\mathbf{x},t) &= -u(\mathbf{x},t) + \alpha(\mathbf{x},t)S(\mathbf{x},t) \\ &\quad + \beta \int w(\mathbf{x}-\mathbf{x}')' f[u(\mathbf{x}',t)]d\mathbf{x}' + \gamma\sigma(\mathbf{x},t) + h \\ \text{where } f[u(\mathbf{x},t)] &= \frac{1}{1+\exp(\frac{-2(u(\mathbf{x},t)-\theta(\mathbf{x},t))}{\nu(\mathbf{x},t)})} \end{aligned} \tag{1}$$

Here, the quantity $u(\mathbf{x},t)$ represents the membrane potential of the field at time t and position $\mathbf{x}$, $S(\mathbf{x},t)$ the afferent input, $w(\mathbf{x}-\mathbf{x}')$ the fixed lateral *interaction kernel*, $f[u]$ the nonlinearity or *transfer function*, and $\sigma(\mathbf{x},t)$ the noise. τ determines the time scale of field evolution, and h is the *resting potential*, i.e., the equilibrium potential in case of no input. We choose a sigmoid transfer function, parameterized for each neuron by a threshold and a gain value $\theta(\mathbf{x},t), \nu(\mathbf{x},t)$. In addition to the original model equation [8], we introduce time and position dependent coefficients $\alpha(\mathbf{x},t)$, $\theta(\mathbf{x},t)$, $\nu(\mathbf{x},t)$, as well as the coefficients β, γ which will not be subject to adaption for now. The coefficients $\alpha(\mathbf{x},t), \beta$ and γ respectively determine the contribution of the afferent input, the lateral recurrent interactions and the noise. The interaction kernel $w(\mathbf{x}-\mathbf{x}')$ is usually chosen to be symmetric: $w(\mathbf{x}-\mathbf{x}') = a_0 G_{\mu=0,\sigma_{\text{on}}}(\mathbf{x}-\mathbf{x}') - b_0 G_{\mu=0,\sigma_{\text{off}}}(\mathbf{x}-\mathbf{x}') - c_0$, where $G_{\mu=0,\sigma}(\mathbf{x})$ denotes a Gaussian with mean μ and standard deviation σ, and $\sigma_{\text{on}} < \sigma_{\text{off}}$. The constants a_0, b_0, c_0 are chosen suitably to achieve the desired level of local excitation/inhibition (a_0, b_0) as well as global inhibition (c_0).

3 Experimental Setup

The configuration used for simulation experiments consists of a single neural field discretized to 128x128 neurons. Constant parameters are chosen to $\tau = 12$, $a_0 = 0.3, b_0 = 1.5, \sigma_{\text{on}} = 10, \sigma_{\text{off}} = 20, \beta = 1, \gamma = 0, h = -0.15$. Input patterns stay constant for one *pattern cycle* consisting of 800 iteration steps. For clearing activity from previous pattern cycles, a value of $h = -20.0$ is used in the first 150 steps of a pattern cycle. Afferent input $S(\mathbf{x},t)$ is additively composed of uniform noise with amplitudes between 0.14 and 0.16 and Gaussians (random peak values between 0.5 and 0.7, random variances between 4.0 and 6.0) that appear, at the start of each pattern cycle, at random positions in two distinct 30x30 areas (see Fig. 1).

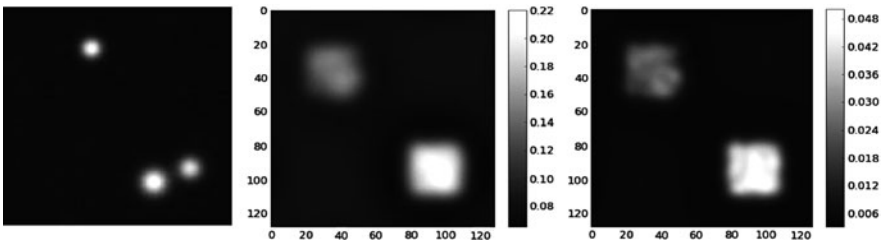


Fig. 1 *Left*: Example input stimulus. *Middle,right*: Temporal mean and variance of membrane potentials $u(\mathbf{x},t)$ resulting from input stimuli, without adaptation

One area always contains two Gaussians whereas an other area contains only one. Thus, we can distinguish three activity levels in the afferent input. Using fixed values of $\alpha(\mathbf{x}, t) = 1.0$, $\nu(\mathbf{x}, t) = 0.3$ and $\theta(\mathbf{x}, t) = 0.5$, the mean and variance distribution resulting from this input can be seen in Fig. 1.

4 Dynamic Adaptation of Membrane Potential

In order to regulate the temporal mean value of the membrane potential $\bar{u}^{\lambda}(\mathbf{x}, t)$ to a target value $\bar{u}_{\text{target}}$, we use a gradient adaptation rule for the input strength $\alpha(\mathbf{x}, t)$ of each neuron:

$$\alpha(\mathbf{x}, t+1) = \alpha(\mathbf{x}, t) - \epsilon_{\alpha}(\bar{u}^{\lambda}(\mathbf{x}, t) - \bar{u}_{\text{target}}). \tag{2}$$

$$\bar{u}^{\lambda}(\mathbf{x}, t+1) = (1-\lambda)\bar{u}^{\lambda}(\mathbf{x}, t) + \lambda u(\mathbf{x}, t). \tag{3}$$

Here, λ and ϵ_{α} denote the timescales at which mean calculation and adaptation take place. Importantly, λ and ϵ_{α} have to be properly chosen. If the adaptation is too fast, i.e. ϵ_{α} is too large, it has an immediate effect on the field potential, but $\alpha(\mathbf{x}, t)$ will vary constantly and not stabilize. If the adaptation is slow, $\alpha(\mathbf{x}, t)$ will converge slowly without oscillating. An aim of this section is to find a suitable combination (ϵ_{α},λ) that guarantees stable convergence of $\alpha(\mathbf{x}, t)$.

4.1 Experiments and Results

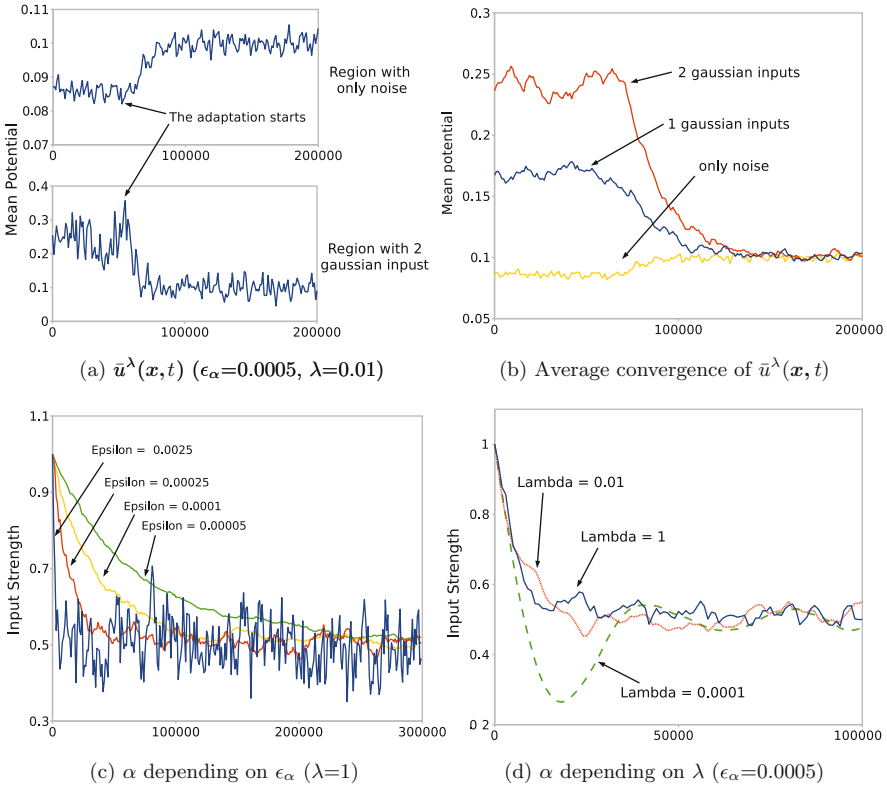
Considering the distribution of the mean potential without adaptation (Fig. 1), we set the target mean potential $\bar{u}_{\text{target}}$ to 0.1 and study the convergence of the mean potential depending on ϵ_{α} and on λ.

The results in Fig. 2 show that we can successfully adapt the mean potential using the gradient adaptation rule. If the adaptation is on a faster time scale than the mean calculation, the input strength will oscillate, as can be observed in Fig. 2d with λ=0.0001. Therefore, a general rule for input strength adaptation is to set different time scales satisfying $\lambda \gg \epsilon_{\alpha}$.

5 Dynamic Adaptation of Firing Rates

This section describes how to adapt the parameters $\theta(\mathbf{x}, t)$ and $\nu(\mathbf{x}, t)$ of the transfer function in order to control the mean and the standard deviation $\Sigma_f^{\rho}(\mathbf{x}, t)$ of each neuron's firing rate. Mathematically, we can express a running estimate on time scale ρ of the firing rate's mean and standard deviation as:

$$\bar{f}^{\rho}(\mathbf{x}, t+1) = (1-\rho)\bar{f}^{\rho}(\mathbf{x}, t) + \rho f(\mathbf{x}, t) \tag{4}$$



(a) $\bar{u}^{\lambda}(\boldsymbol{x},t)$ (ϵ_{α}=0.0005, λ=0.01)

(b) Average convergence of $\bar{u}^{\lambda}(\boldsymbol{x},t)$

(c) α depending on ϵ_{α} (λ=1)

(d) α depending on λ (ϵ_{α}=0.0005)

Fig. 2 Mean potential adaptation. (**a**) shows the convergence of the mean potential $\bar{u}^{\lambda}(\mathbf{x},t)$. (**b**) shows the average convergence of $\bar{u}^{\lambda}(\mathbf{x},t)$ for a set of 10 experiments, with the same parameters as in (**a**). Fig. (**c**) and (**d**) show the convergence of $\alpha(\mathbf{x},t)$ depending on ϵ_{α} and λ, in region with 2 gaussian inputs

$$\Sigma_f^{\rho}(\mathbf{x},t+1) = (1-\rho)\Sigma_f^{\rho}(\mathbf{x},t) + \rho\sqrt{(f[u(\mathbf{x},t)] - \bar{f}^{\rho}(\mathbf{x},t))^2} \tag{5}$$

For the dynamic firing rate adaptation, we adapt the threshold $\theta(\mathbf{x},t)$ to the mean value of the potential:

$$\theta(\mathbf{x},t+1) = \theta(\mathbf{x},t) - \epsilon_{\theta}(\theta(\mathbf{x},t) - \bar{u}^{\lambda}(\mathbf{x},t)) \tag{6}$$

By doing so, we center the transfer function on the mean value of the potential distribution. This is essential if we want to adapt the gain efficiently. Next, as illustrated in Fig. 3, we adjust the gain in order to adapt $\Sigma_f^{\rho}(\mathbf{x},t)$:

$$\nu(\mathbf{x},t+1) = \nu(\mathbf{x},t) - \epsilon_{\nu}(\Sigma_f^{\rho}(\mathbf{x},t) - \Sigma_{\text{target}}) \tag{7}$$

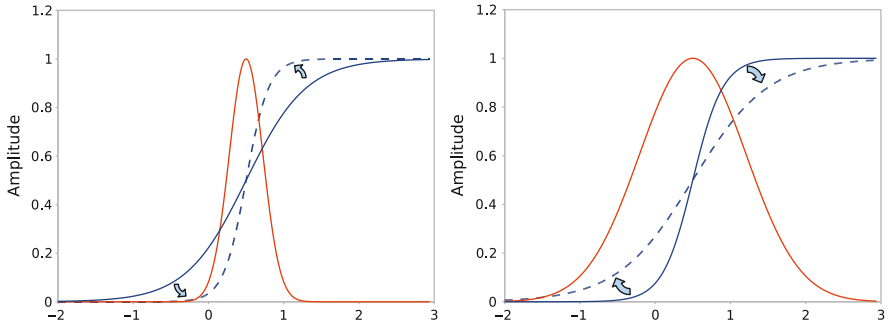


Fig. 3 Mechanisms of gain adaptation. *Red solid curves* indicate the distribution of membrane potential values, *blue solid* and *dashed curves* indicate the current and the optimal transfer function. As the gain is inversely proportial to the slope of the sigmoid function, increasing (*left*) or decreasing (*right*) the gain will both decrease or increase the firing rate variance and adapt the transfer function (*blue arrows*) so that it maps most potential values linearly (Colors will be viewed in online version)

5.1 Experiments and Results

For simulations, we use the following values: $\Sigma_{\text{target}} = 0.015$, $\rho = 0.01$, $\epsilon_\theta = 0.0001$ and $\epsilon_\nu = 0.00001$. As one can see in Fig. 4b, the transfer function adaptation is efficient for areas with strong input (2 gaussian inputs). However, for the case of a region with a weak input (only noise), the activity may die out due to lateral interactions (see Fig. 4a) . This can be avoided by combining the firing rate adaptation with the mean potential adaptation (see Section 6)

6 Combination of Dynamic Adaptation Methods

In this section, it will be described how the self-adaptation of the potentials and the firing rates is performed simultaneously. In order to avoid interferences, attention must be given to the time scales of the adaptation mechanisms ρ and λ, and to the adaptation constants ϵ_α, ϵ_ν and ϵ_θ.

We require that all statistical quantities should operate on a similar time scale ($\lambda \approx \rho$), and that adaptation mechanisms should be significantly slower. Threshold adaptation is coupled to input strength adaptation, then $\epsilon_\alpha \approx \epsilon_\theta$. Gain adaptation depends critically on centering of the threshold, therefore $\epsilon_\nu < \epsilon_\theta$:

$$1/\tau > 1/800 > \lambda \approx \rho > \epsilon_\alpha \approx \epsilon_\theta > \epsilon_\nu \tag{8}$$

Using target values of $\bar{u}_{\text{target}}$=0.1, Σ_{target}=0.015, time constants λ=0.01 and ρ=0.01, and adaptation coefficients ϵ_α=0.0005, ϵ_θ=0.0001, and ϵ_ν=0.00001, one can see in Fig. 5 that the joint adaptation is indeed successful. We do not observe the previous problem with weak input strength anymore.

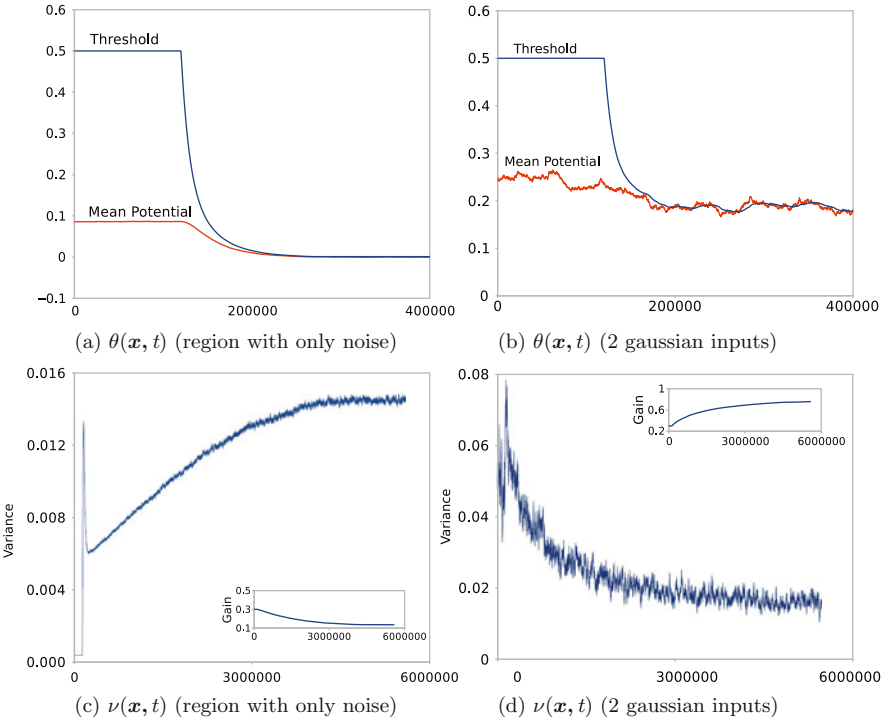
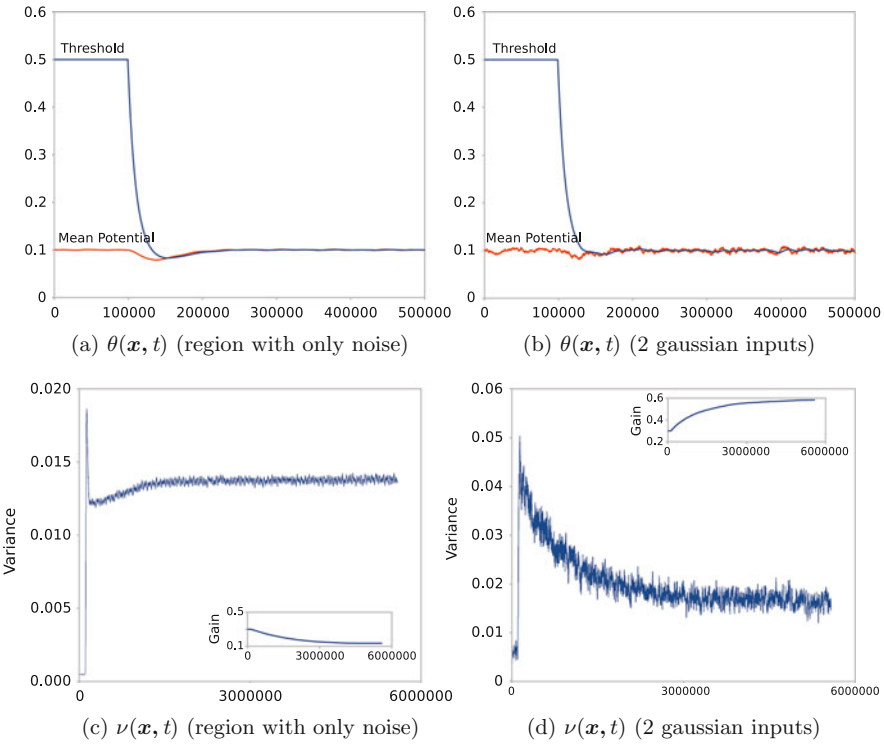


Fig. 4 Transfer function adaptation. The gain and the threshold both converge and lead to a stable system in the case of a strong input (**b, d**). The mean potential converge to 0 for a low input (**a, c**). Variance plots are obtained by averaging a set of 10 experiments

7 Discussion

This contribution introduced several adaptation rules for actively maintaining neurons in a desired dynamic state by suitably modifying their internal processing parameters. An important outcome of our experiments is the fact that such adaptation mechanisms must work on appropriate time scales w.r.t. each other in order to avoid instabilities and divergences. The benefit of the work presented here can be summarized as follows: First of all, the number of free parameters that have to be set by a system designer is reduced. Secondly, neural fields can now be used in a "plug and play" manner, connecting them without having to explicitly consider parameter settings for dynamical stability. This will facilitate the construction of large neural systems as envisioned in [23, 7]

However, there still remains a significant number of parameters that are not subject to adaptation. Many of these parameters may be eliminated by further adaptation mechanisms, which will be part of our future work. However, there must remain a set of parameters describing some high-level function of a neuron (e.g., accumulation, decision making, integration, discrimination) that are determined by a designer.



(a) $\theta(\boldsymbol{x}, t)$ (region with only noise) (b) $\theta(\boldsymbol{x}, t)$ (2 gaussian inputs)

(c) $\nu(\boldsymbol{x}, t)$ (region with only noise) (d) $\nu(\boldsymbol{x}, t)$ (2 gaussian inputs)

Fig. 5 Result with both mean potential adaptation and transfer function adaptation. Variance plots are obtained by averaging a set of 10 experiments

Therefore it should be clear that self-adaptation can eliminate several but not all free parameters in a system.

8 Conclusion

As was mentioned before, there exists a number of parameters that is not considered for adaptation in this contribution. In the future, it should be investigated how to extend self-adaptation mechanisms to these parameters (which are treated as constants right now), most notably the lateral interaction strength β and the time constant τ.

References

1. Izhikevich, E.M.: Which model to use for cortical spiking neurons? IEEE Trans.Neural Netw. **15**(5) (2004) 1063–1070.
2. Turrigiano, G.G., Nelson, S.B.: Homeostatic plasticity in the developing nervous system. Nat. Rev. Neurosci. **5**(2) (Feb 2004) 97–107.

3. Turrigiano, G.: The self-tuning neuron: synaptic scaling of excitatory synapses. Cell **135**(3) 2008.
4. Gepperth, A., Fritsch, J., Goerick, C.: Cross-module learning as a first step towards a cognitive system concept. In Proceedings of the First International Conference on Cognitive Systems. IEEE Press (2008).
5. Fay, R., Kaufmann, U., Knoblauch, A., Markert, H., Palm, G.: Combining visual attention, object recognition and associative information processing in a neurobotic system. In Wermter, S., Palm, G., Elshaw, M.,eds.: Biomimetic Neural Learning for Intelligent Robots. Intelligent Systems, Cognitive Robotics, and Neuro-science., vol. 3575 of Lecture Notes in Computer Science LNAI. Berlin Heidelberg: Springer (2005) 118–143.
6. Wennekers, T., Garagnani, M., Pulvermüller, F.: Language models based on hebbian cell assemblies. J. Physiol. Paris **100**(1–3) (2006) 16–30.
7. Vernon, D., Metta, G., Sandini, G.: The iCub cognitive architecture: Interactive development in a humanoid robot. In 6th International Conference on Development and Learning (2007).
8. Amari, S-I.: Mathematical foundations of neurocomputing. Proc. IEEE **78**(9) (1990) 1441–1463.
9. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. Psychol. Rev. **109**(3) (Jul 2002) 545–572.
10. Cisek, P.: Integrated neural processes for defining potential actions and deciding between them: a computational model. J. Neurosci. **26**(38) (Sep 2006) 9761–9770.
11. Zhang, W., Linden, D.J.: The other side of the engram: experience-driven changes in neuronal intrinsic excitability. Nat. Rev. Neurosci. **4** (2003) 885–900.
12. Stemmler, M., Koch, C.: How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. Nat. Neurosci. **2**(6) (1999) 521–527.
13. Triesch, J.: Synergies between intrinsic and synaptic plasticity mechanisms. Neural Comput. **19**(4) (2007) 885–909.
14. Elliott, T., Kuang, X., Shadbolt, N.R., Zauner, K.P.: An invariance principle for maintaining the operating point of a neuron. Network **19**(3) (2008) 213–235.
15. Steil, J.J.: Online reservoir adaptation by intrinsic plasticity for backpropagationdecorrelation and echo state learning. Neural Netw. **20**(3) (April 2007) 353–364.
16. Butko, N.J., Triesch, J.: Learning sensory representations with intrinsic plasticity. volume 70, pp. 1130–1138, 2007. Advances in Computational Intelligence and Learning – 14th European Symposium on Artificial Neural Networks (2006).
17. Gläser, C., Joublin, F., Goerick, C.: Homeostatic development of dynamic neural fields. In International Conference on Development and Learning, pp. 121–126 (2008).
18. Johnson, J.S., Spencer, J.P., Schöner, G.: Moving to higher ground: the dynamic field theory and the dynamics of visual cognition. New Ideas Psychol. **26**(2) (Aug 2008) 227–251.
19. Wilimzig, C., Schneider, S., Schöner, G.: The time course of saccadic decision making: dynamic field theory. Neural Netw. **19**(8) (Oct 2006) 1059–1074.
20. Rougier, N.P., Vitay, J.: Emergence of attention within a neural population. Neural Netw. **19**(5) (Jun 2006) 573–581.
21. Faubel, C., Schöner, G.: Fast learning to recognize objects: dynamic fields in labelfeature spaces. In Proceedings of the 5th International Conference on Development and Learning (2006).
22. Deco, G., Rolls, E.T.: A neurodynamical cortical model of visual attention and invariant object recognition. Vision Res. **44**(6) (Mar 2004) 621–642.
23. Gepperth, A., Fritsch, J., Goerick, C.: Computationally efficient neural field dynamics. In Proceedings of the 16th European Symposium on Artificial Neural Networks, pp. 179–185 (2008).

Functions of Shallow vs. Deep Theta-Nested Gamma in Object Recognition

Xuejuan Zhang, Yang Zhan, and Keith M. Kendrick

Abstract Using a network of excitatory pyramidal cells and inhibitory interneurons with fast and slow $GABA_A$ kinetics, we have successfully reproduced our recent experimental observations from local field potential and multiunit activity recordings from the sheep's inferior temporal cortex during learning of visual discrimination tasks. Both experimental and numerical results showed that following learning, theta wave amplitude is significantly increased and its phase is more highly correlated with the amplitude of gamma wave during stimulus presentation. The increased correlation between the theta phase and the gamma amplitude is a consequence of theta-nested gamma. Importantly, we have used the model to show that learning effects can be reproduced simply by altering the sensitivity of NMDA receptors and that the resultant shallowing of theta-nested gamma potentiates responses made by downstream neurons.

Keywords Integrate-and-fire neural network · Theta-nested gamma rhythm · AMPA receptors · $GABA_A$ receptors · Time-frequency spectrum

1 Introduction

Rhythmic, oscillatory activity in the brain is believed to play an important role in the control of a number of different behavioral states in humans and other animal species. In the literature, the most frequently studied brain regions are the hippocampus and neocortex, where different frequency activities are associated with higher cognitive functions such as perception, categorization and memory formation, and slow and frequency waves usually occur simultaneously in brain information processing [1]. With the progress of new recording techniques and the development of mathematical tools such as wavelet analysis and causal analysis [2–4], several

X. Zhang (✉)
School of Mathematical and Physical Sciences, Zhejiang Normal University, Jinhua 321004, P.R. China
e-mail: xuejuanzhang@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_84,

cross-frequency interactions were revealed, e.g., n: m amplitude-independent phase coupling [5], and the modulation of the amplitude of fast rhythm by the phase of the slow frequency wave [6–8], etc. Recently, it was proposed that the most readily observable form of cross-frequency interaction is that of "nesting" of the higher frequency oscillation on the lower frequency one [7]. Though there is extensive experimental support for this form of "nesting" from the hippocampus and neocortex, the exact neuronal mechanism and the functional role of such nested oscillations in brain information processing still remains unclear.

In the present work, we first analyzed 64-channel multi-electrode data recorded in sheep's inferior temporal (IT) cortex while it performed visual discriminations between pairs of faces or other objects. It was found that there was significant coupling between theta (4–8 Hz) and gamma (30–80 Hz) waves in >75% of electrodes and that the strength of this coupling increased during stimulus presentation as a function of learning. To understand the mechanism of the theta-nested gamma rhythm in object recognition, we then constructed a network composed of three populations of integrate-and-fire neurons based on two kinetically distinct $GABA_A$ receptor-mediated currents [9–11] to reproduce corresponding phenomena observed in the experiment. Our model suggests that to perform discrimination task successfully, gamma rhythm should be shallowly but not deeply nested in theta wave.

2 Material and Methods

2.1 Subject and Experimental Recordings

In the experiment, after operant visual discrimination training, three sheep were implanted with chronic 64-channel multi-electrode arrays (MEAs – tungsten wires with 250 μm spacing) in the right (n = 3) and left (n = 2) IT cortices under general (fluothane) anaesthesia and with full sterile precautions. Following recovery, recordings were made from the unanaesthetised sheep while they performed operant discriminations between different pairs of sheep faces or object. A correct panel-press choice response elicited a food reward (see Fig. 1a). For the recordings the MEAs were connected to a 128 channel recording system (Black Rock Microsystems, USA). All experimental procedures involving animals were conducted in strict accordance with The Animals Scientific Procedures Act 1986, UK.

Local field potentials (LFPs) were recorded simultaneously from 128 chronically implanted electrodes. The LFPs were sampled at 2,000 Hz and digitized for storage from around 3 s prior to the stimulus onset to 3 s after the stimulus onset. The data was from 4 to 10 different stimulus pairs in each animal with recording sessions being carried out on repeated occasions where the animals were given 20–40 trial blocks for each stimulus pair.

2.2 Model

We constructed a network consisting of three populations of neurons: 100 excitatory (pyramidal) neurons, 50 inhibitory fast (inter) neurons and 50 inhibitory slow (inter)

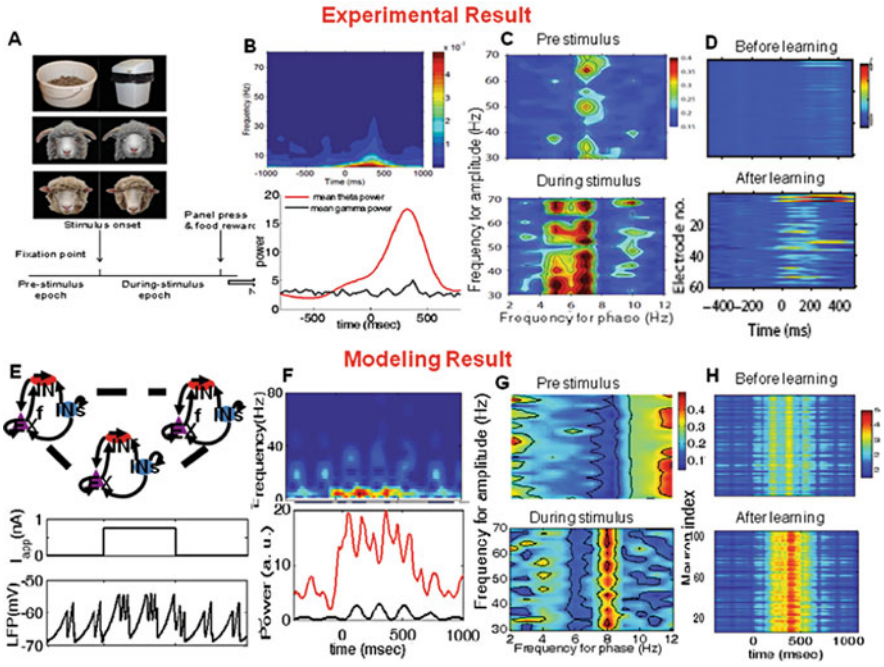


Fig. 1 **a–d** *Experimental results*: **a** Examples of different pairs of sheep faces or objects the sheep were required to discriminate between to obtain a food reward. **b** Trial-averaged wavelet power. Time-dependent mean powers in the theta band (*red curve*) and in the gamma band (*black curve)* are shown in the bottom panel. **c** Cross-frequency coupling between gamma amplitude and theta phase for the pre-stimulus epoch (*the upper panel*) and during-stimulus epoch (*the bottom panel*) following learning. **d** Theta/gamma ratio across IT electrodes before (*the upper panel*) and after learning (*the bottom panel*). There is an increase during the stimulus after learning. **e–g** *Computational results*: **e** A network of 100 excitatory neurons (EX) and 50 fast inhibitory neurons (INf) and 50 slow inhibitory neurons (Ins). Only three of each type of neuron are plotted here to show connections. **f–h** are from model simulations corresponding to **b–d** in the IT experiments respectively

neurons with all-to-all connections (see Fig. 1 g). Each set of neurons obeys an integrate-and-fire equation:

$$C\frac{dV(t)}{dt} = -g_L(V - E_L) - I_{syn} + I_{app}, \tag{1}$$

where C is the capacitance for the neuron, g_L is the leaky conductance, I_{syn} is the synaptic input from other neurons and I_{app} is the external input. When V reaches a firing threshold V_{th} a spike is discharged and $V$is reset to V_{rest} and stays there for an absolute refractory periodτ_{ref}. For excitatory neurons, we set $C = 0.5$nF, $g_L = 0.025\,\mu$s, $E_L = -70$mV, $V_{\text{th}} = -52$ mV, Vrest$=-59$ mV£, $\tau_{ref} = 2$ ms £ while for inhibitory neurons, we set $C = 0.2$nF, $g_L = 0.02\,\mu$s, $E_L = -65$ mV £, $V_{\text{th}} = -52$ mV, $V_{rest} = -60$ mV, $\tau_{ref} = 1$ ms.

Each neuron receives AMPA and NMDA receptor-mediated currents from excitatory (EX) cells, GABA_A receptor mediated currents from fast inhibitory (INf)

neurons as well as slow inhibitory (INs) neurons. The gating variable s for AMPA and NMDA receptors is described by two first-order kinetics [12]:

$$\frac{dx}{dt} = \alpha_x \sum_j \delta(t - t_j) - x/\tau_x, \quad \frac{ds}{dt} = \alpha_s x(1 - s) - s/\tau_s, \tag{2}$$

where t_j is the presynaptic spike time. We used $\alpha_x = 1$(in dimensionless), $\tau_x = 0.05\,\text{ms£}$, $\alpha_s = 1.0\,\text{ms}^{-1}\text{£}$, $\tau_s = 2.0\,\text{ms}$ for AMPA receptors, and $\alpha_x = 1$ (in dimensionless), $\tau_x = 2\,\text{ms£}$, $\alpha_s = 1.0\,\text{ms}^{-1}\text{£}$, $\tau_s = 80\,\text{ms}$ for NMDA receptors. The gating variable s_{GABA} for GABA_A receptors obeys a simple first-order kinetics [13]:

$$\frac{ds_{GABA}}{dt} = \alpha_I \sum_j \delta(t - t_j^-)(1 - s_{GABA}) - s_{GABA}/\tau_I, \tag{3}$$

where t_j^- indicates the time immediately before the spike at time t_j. We used $\tau_I = 9\,\text{ms}, \alpha_I = 1\,\text{ms}^{-1}$ for the fast GABA_A channels, and $\tau_I = 50\,\text{ms}, \alpha_I = 0.2\,\text{ms}^{-1}$ for the slow GABA_A channels. Thus the AMPA and NMDA receptors-mediated currents are given by: $I_{AMPA} = g_{AMPA} s_{AMPA}(V - V_E)$, and $I_{NMDA} = g_{NMDA} s_{NMDA}(V - V_E)$, respectively, and the GABA_A receptor-mediated current is given by $I_{GABA} = g_{GABA} s_{GABA}(V - V_I)$. Here $V_E = 0\,\text{mV}$, $V_I = -70\,\text{mV}$.

We assumed that all neurons receive background currents all of the time. These were set as: 0.7(1±10%) nA for EX neurons, 0.85 nA for INf neurons and 0.6 nA for INs neurons. The stimulus was assumed to be added only to the EX neurons. Other parameters are set as: $g_{AMee} = 0.02$, $g_{AMef} = 0.08$, $g_{AMes} = 0.0005$, $g_{GAfe} = 0.01$, $g_{GAff} = 0.08$, $g_{GAfs} = 0.0$, $g_{GAss} = 0.08$, $g_{GAse} = 0.06$, $g_{GAsf} = 0.03$, $g_{NMee} = 0.002$, $g_{NMes} = 0.0001$, $g_{NMef} = 0.001$, except $g_{GAfe} = 0.007$ for shallow nested gamma and $g_{GAfe} = 0.02$ for deep nested gamma in Fig. 2; and $g_{NMee} = 0.0035$, $g_{NMes} = 0.00055$in the bottom panel of Fig. 1 h. Here AM refers to AMPA receptors, NM refers to NMDA receptors, and GA refers to GABA_A receptors, *ee* refers to recurrent connection among excitatory neurons, *ef* refers to EX→INf connection, *es* refers to EX→INs connection, and so on. The amplitude of the stimulus was set as $I_{Amp} = 0.8\,\text{nA}$ except Fig. 2d, e.

3 Results

3.1 Computational vs. Experimental Results

To test how reliable the constructed neuronal network is, we have to implement the model to reproduce the phenomena observed in the experiment. Let us first briefly summarize the sheep IT data, for detailed analysis, see [14]. Using local field potential and multi-unit neuronal activity (MUA) recording from 64-electrode arrays in sheep IT cortex, it was found that an evoked response in both MUA and visual evoked potential occurred at about 300 ms after the presentation of stimulus

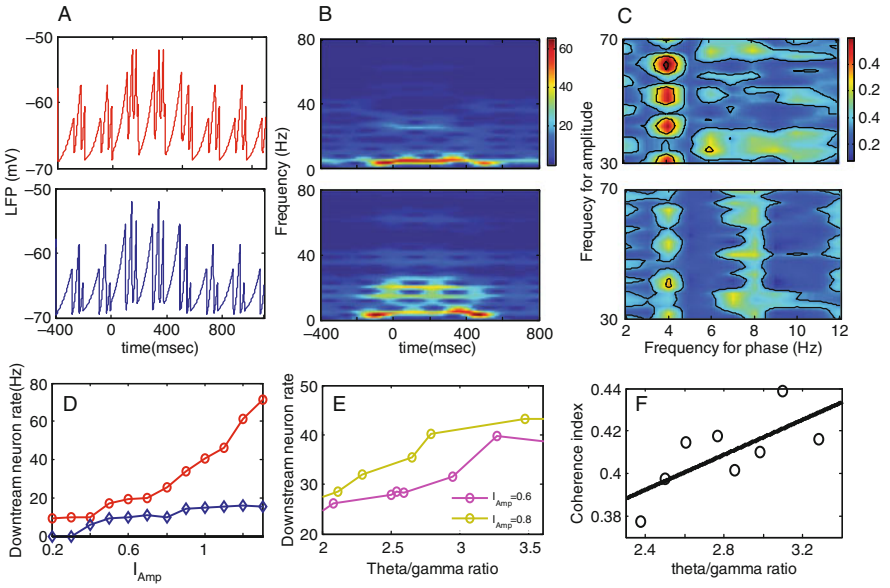


Fig. 2 *Functional significance of shallow theta-nested gamma* **a** Shallow (*top*) and deep (*bottom*) theta-nested gamma oscillations typical of what is seen in IT after and during learning respectively. **b–c** Time-dependent power spectra and the cross-frequency coherence corresponding oscillations in **a**. **d** The firing rates of a downstream neuron at different stimulus intensities for shallow (*red*) and deep (*blue*) theta-nested gamma rhythms. **e** The firing rate of a downstream neuron vs. the theta/gamma ratio. **f** Cross-frequency coherence vs. the theta/gamma ratio

(Fig. 1b), and that following learning the amplitude of theta, but not gamma band oscillations increased at this same time point (data not shown here, see [14]). There was coupling between theta phase and gamma amplitude across >75% of electrodes, but following learning such a cross-frequency coupling was much stronger in during the stimulus period than in pre-stimulus period (Fig. 1c) and the ratio of theta amplitude vs. gamma amplitude was also increased (Fig. 1d). Furthermore, the magnitude of such changes correlated significantly with discrimination performance. There was no learning evoked change in the MUA however which is consistent with the lack of change in gamma amplitude.

To compare the model with experimental results, Fig. 1e–h gave the corresponding computational results. The LFP averaged over 100 Excitatory (EX) neurons in response to a square-shaped stimulus was plotted in the bottom panel of Fig. 1e, where the averaged LFP exhibits theta-nested gamma rhythm and there is increased theta, but not gamma amplitude at around 300 ms after stimulus onset. Figure 1f is the averaged time-dependent power spectrum over the 100 EX neurons. To provide a quantitative comparison, the time-dependent power in both theta and gamma bands were averaged. In Fig. 1f it can be seen that at around 300–350 ms theta power reaches its maximum and is about 10 times higher than that of gamma. This result is consistent with the experimental observations in IT shown in Fig. 1b.

What about the cross-frequency coupling between the gamma amplitude and theta phase in the modeling network? The patterns of coherence between these two frequencies before and during stimulus presentation were shown in Fig. 1 g. It is known that altered NMDA receptor sensitivity underlies many forms of neural plasticity changes associated with learning. Our computational result also showed that by properly strengthening the EX→EX and EX→INs connections mediated via NMDA receptors, the changes in theta-nested gamma seen following learning in the IT could be produced, i.e., the theta/gamma ratio is increased (see Fig. 1 h). This is also consistent with the experimental recordings in IT.

3.2 Functions of Shallow Theta-Nested Gamma Rhythm

We next used the model to investigate in more detail the functional consequences of the shift from deep to shallow nested gamma seen following learning in the IT. When we look at the power spectrum and the coherence of the coupling between theta phase and gamma amplitude for the shallow and deep gamma, respectively, Fig. 2a–c show that when the gamma wave is deeply nested in the theta wave, the power spectrum in the presence of stimulus does not concentrate in the theta band, also the cross-frequency coupling between the theta phase and the gamma amplitude only has a weak coherence. However, for shallow theta-nested gamma oscillation, the power spectrum in the theta band is significantly increased when the stimulus is turned on. Moreover, the coherence between the gamma magnitude and the theta phase is greatly increased. This implies that the shallow theta-nested gamma is necessary for enhancing the cross-frequency coupling between the theta and gamma oscillations.

Next we explored how the depth of the nested gamma oscillation affects the firing behavior of a downstream neuron to which the EX neurons project. In Fig. 2d it can be seen that, for the shallow-nested gamma, the firing rate monotonically increases with the increase of stimulus strength; while for the deep-nested gamma, the curves are more flat with the firing rates below 30 Hz. This shows that the presence of shallow theta-nested gamma can act to potentiate the firing of a downstream neuron. To quantify the effect of the shape of theta-nested gamma rhythm, we calculated the average gamma and theta amplitudes and plotted the firing of the downstream neuron as a function of the size of theta/gamma ratio. In Fig. 2e it can be seen that with the increase of theta/gamma ratio the firing rate of the downstream neuron increases correspondingly.

Finally we explored if the learning-evoked change in the cross-frequency coupling is also the functional consequence of an increased proportion of theta to gamma (Fig. 1 h), i.e., shallow theta-nested gamma. Figure 2f confirms that with the increase of theta/gamma ratio, the coherence of the cross-frequency coupling between the two rhythms increases correspondingly.

4 Discussion

From our model, we have observed widespread increased theta activity during an applied stimulus, similar to that observed from the LFP data in sheep IT cortex during performance of a discrimination task [14]. Furthermore, the phase of the theta wave has been shown to be correlated with the amplitude of the low gamma wave (30–70 Hz), as a consequence of the theta nested-gamma activity. Our computational results together with the experimental observation in IT cortex [14], as well as previous evidence from studies in the hippocampus [6–8]), suggest that shallowing of theta-nested gamma following learning may serve as an effective mechanism for modulating neuronal outputs so that they produce potentiated responses in downstream neurons.

In the current computational study, we only incorporated classic neurotransmitter-mediated currents to illustrate functions of rhythmic oscillations in object recognition and learning. It is known that in all brain areas, neurons also use a huge diversity of slower analog signaling mechanisms, these chemical signaling pathways, acting in a more global spatial scale and on a longer temporal scale, are closely related with coherent behavior including learning and memory [15, 16]. How do these parallel signaling systems interact with neurotransmitter signals to give rise to long last coherent change in behavior is our next goal of study.

Acknowledgments This work was supported by the National Natural Science Foundation of China (under Grant Nos. 10771155, 10971196), and a Foundation for the Author of National Excellent Doctoral Dissertation of PR China (FANEDD).

References

1. Buzsäki, G.: Rhythms of the Brain. Oxford: Oxford University Press (2006).
2. Wu, J.H., Liu, X., Feng, J.: Detecting causality between different frequencies. J. Neurosci. Meth. **167** (2008) 367–375.
3. Guo, S., Wu, J., Ding, M., Feng, J.: Uncovering interactions in the frequency domain. PLoS Comput. Biol. **4** (2008) e1000087.
4. Zou, C., Feng, J.: Granger causality vs. Dynamic Bayesian network inference: a comparative study. BMC Bioinform. **10** (2009) 122 doi:10.1186/1471-2105-10-122.
5. Palva, J.M., Palva, S., Kaila, K.: Phase synchrony among neuronal oscillations in the human cortex. J. Neurosci. **25** (2005) 3962–3972.
6. Lakatos, P., et al.: An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J. Neurophys. **94**(3) (2005) 1904–1911.
7. Roopun, A.K., Kramer, M.A., et. al.: Temporal interactions between cortical rhythms. Front. Neurosci. **2** (2008) 145.
8. Canolty, R.T., et al.: High gamma power is phase-locked to theta oscillations in human neocortex. Science **313** (2006) 1626–1628.
9. Banks, M.I., Li, T.B, Pearce, R.A: The synaptic basis of gabaa,slow. J. Neurosci. **18** (1998) 1305–1317.
10. Sceniak, M.P, MacIver, M.B.: Slow GABAA mediated synaptic transmission in rat visual cortex. BMC Neurosci. **16** (2008) doi: 10.1186/1471-2202-9-8.

11. White, J., Banks, M., Pearce, R., Kopell, N.: Networks of interneurons with fast and slow GABAA kinetics provide substrate for mixed gamma-theta rhythm. PNAS **97** (2000) 8128–8133.
12. Wang, X-J.: Synaptic basis of cortical persistent activity: the importance of NMDA receptors toworking memory. J. Neurosci. **19** (1999) 9587–9603.
13. Wang, X-J, Rinzel, J.: Alternating and synchronous rhythms in reciprocally inhibitory model neurons. Neural Comput. **4** (1992) 84–97.
14. Kendrick, K. M., et. al.: Learning alters Theta-Nested Gamma Oscillations in inferotemporal cortex. Nat. Precedings (2009) hdl: 10101/ npre. 2009.3151.1.
15. Rossoni, E., et. al.: Emergent synchronous bursting of oxytocin neuronal network. PLoS Comput. Biol. **4** (2008) e1000123.
16. Lee, H. J., et al.: Oxytocin: the great facilitator of life. Prog. Neurobiol. **88** (2009) 127–151.

Part IX
Multiscalar Neurodynamics – From Physiology to Systems Theory

Modelling the Hypothalamic Control of Thalamic Synchronization Along the Sleep-Wake Cycles

Svetlana Postnova, Karlheinz Voigt, and Hans Albert Braun

Abstract Thalamic activity significantly changes at the transitions form conscious wake to unconscious sleep states, going from an unsynchronized single-spike firing to synchronized discharges of impulse groups. These transitions can be simulated with a conductance-based model of neurons and synapses which assumes that alterations of thalamic synchronizations results from the neurons′ intrinsic dynamics which, however, are externally tuned by synaptic input from sleep-wake controlling areas in the hypothalamus. The crucial role is assigned to activity dependent changes of the synaptic efficacy of the neuropeptide hypocretin/orexin.

Keywords Neuronal synchronization · Burst discharges · Sleep-wake cycles · Hypocretin/orexin · Synaptic plasticity

1 Introduction

Conscious perception seems to be intimately related to the activity state of the thalamus which is considered to be a gate of sensory information transmission to the cortex [1, 2]. The most significant changes of the gating properties appear at the transitions from wakefulness to sleep when conscious perception is switched off and sensory information transmission is almost blocked. During sleep, only comparably strong stimuli can pass.

Sleep-wake transitions are clearly correlated with distinct changes of thalamic impulse patterns and synchronization states, see e.g. [3, 4]. During sleep, thalamic neurons exhibit synchronized discharges of impulse groups (bursts) while they are in an unsynchronized state of single spike activity (tonic firing) during wakefulness. It

K. Voigt (✉)
Neurodynamics Group, Institute of Physiology, University of Marburg, Marburg D-35037, Germany
e-mail: voigt@staff.uni-marburg.de

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_85,

can be expected that individual neurons which are strongly synchronized with each other cannot react as sensitive on external stimuli as those in an unsynchronized state.

Despite the physiological consequences of thalamic synchronization and its obvious relation to the sleep-wake states, the underlying mechanisms are still under debate. The origin is often assumed in the thalamocortical interactions with ongoing discussions of whether the cortical or the thalamic neurons or their interconnections play the major role, e.g. [1, 2]. Several large-scale models have been presented to simulate the dynamics of these brain areas during the states of sleep and/or wakefulness [5–7]. We are focusing on the transitions between sleep and wake states which seem to be under control of subthalamic, specifically hypothalamic nuclei [8].

According to the generally accepted concept of Borbély, the sleep-wake transitions are determined by the interaction of a circadian and a homeostatic process [9]. An endogenous circadian process has undoubtedly been identified in the genetic clocks of suprachiasmatic hypothalamic neurons (SCN) [10] while the neuronal basis of the homeostatic sleep mechanisms is still purely understood. Nevertheless, it can be expected that the homeostatic mechanisms are related to the activity of hypothalamic nuclei, as hypothalamus is a major centre for regulation of sleep [8, 11].

In our approach, the wake-active neurons in the lateral hypothalamic area (LHA) play the major role. They have excitatory projections to almost all brain areas with glutamate as the main transmitter and the neuropeptide hypocretin (also called orexin) as an obviously most relevant co-transmitter [12–14]. Lack of hypocretin (HCRT) neurons as well as reduced availability of the co-transmitter hypocretin (hcrt) or its postsynaptic receptors leads to the sleep disorder narcolepsy which is characterized by unpredictable transitions between wakefulness and sleep [15, 16].

According to these experimental and clinical data, we have developed a novel concept of homeostatic sleep-wake regulation and have transferred it into a mechanism-based mathematical model which, in connection with a circadian input, can account for periodic transitions between sleep and wakefulness [17]. For the present study, we have combined the model of sleep-wake regulation with a previously developed model of neuronal synchronization which is related to the alterations of impulse pattern [18].

2 Model Structure and Equations

The focus of this modeling study of hypothalamic sleep-wake control of thalamic synchronization is laid on the physiological mechanisms of voltage- and transmitter-gated ion channels. Therefore, a minimal approach of only four neurons is used, as schematically illustrated in Fig. 1. The two hypothalamic neurons have reciprocal excitatory connections. One is representing the hypocretin (HCRT) neurons and the other one local glutamate (GLU) interneurons. Thalamic activity is represented by two gap junction coupled neurons, receiving synaptic input from the hypothalamic

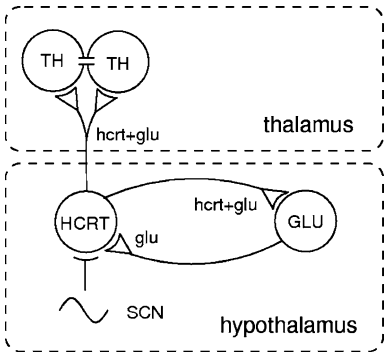


Fig. 1 Schematic representation of the model. The abbreviations are: HCRT – hypocretin neuron, GLU – glutamate neuron, TH – thalamic neurons, SCN – input from the suprachiasmatic nucleus of the hypothalamus glu – glutamate, hcrt –hypocretin, glu+hcrt indicates co-release of the transmitters. For details see text

HCRT neuron. All external influences are comprised in a compound current input to the HCRT neuron. An additional noise current is added to the thalamic neurons to allow de-synchronization out of a synchronized state.

Accordingly, the general form of the membrane equations is given by:

$$CdV/dt = I_l + \sum I_V + \sum I_S + I_{ext} + I_{noise} \tag{1}$$

The changes of the membrane potential dV per time step dt depends on the charge of the membrane capacitance C by a diversity of ion currents I. In all neurons, there is the leak current I_l and there are the fast, voltage dependent Na^+ and K^+ currents for spike generation I_{Na} and I_K. The thalamic neurons additionally include two slowly activating, subthreshold currents $I_{Na,p}$ and $I_{K,Ca}$ for the generation of slow wave potential oscillations and burst discharges. All synaptic currents are modeled with a glutamate term. The synaptic projections from the HCRT neuron include an additional term which accounts for the release of the co-transmitter hcrt.

The gap junction currents between the two neighbored thalamic neurons depend on the conductance of the gap junction g_{gap} and the voltage difference between the neurons:

$$I_{gap} = g_{gap}(V - V_{neighbor}) \tag{2}$$

All other currents I are determined by their maximum conductance $g_{\max}$ multiplied by an activation variable a and the driving force $(V - V_E)$ with V_E as the equilibrium potential:

$$I = g_{\max} a(V - V_E) \tag{3}$$

with $a = 1$ for the leak current with constant conductance.

Voltage-dependent steady state activation a_∞ is modeled in a sigmoid form:

$$a_\infty = 1/(1 + \exp(-s(V - V_H))) \tag{4}$$

with s as the slope and V_H as the half-activation potential.

Activation additionally includes a time delay τ:

$$da/dt = (a - a_\infty)/\tau \tag{5}$$

There are several exceptions for the voltage-dependent currents. For example, the fast sodium current I_{Na} is considered to activate instantaneously with $a = a_\infty$. The slow repolarizing current of thalamic neurons is directly connected to the slow depolarizing current in form of a simplified version of a calcium-dependent potassium current:

$$da_{K,Ca}/dt = (-\eta I_{K,Ca} + k a_{Na,p})/\tau_{K,Ca} \tag{6}$$

where η is the coupling factor and k scales the relaxation.

It is important to note that in case of synaptic currents, although the equations are formally identical with the voltage-dependent currents, the physiological meaning of the parameters and variables can be quite different.

For synaptic currents, the equilibrium potential in Eq. (3) is determined by the ions which the postsynaptic transmitter-gated channels allow to pass. For simplicity, here the V_E is set for both glu- and hcrt-gated ion channels to the *Na* equilibrium potential.

The steady state voltage dependencies of Eq. 4, in contrast, are exclusively determined by presynaptic events. They represent an almost threshold-like curve for transmitter release due to the occurrence of presynaptic spikes.

The connections between presynaptic spikes and postsynaptic currents are provided by the differential equations of the activation variables (Eq. 5). The time constants τ determine the amplitude and delay of postsynaptic currents activations a in response to the presynaptic transmitter release a_∞. With appropriate adjustment, a presynaptic spike can introduce strong but only short postsynaptic currents via ionotropic glu receptors in parallel with weaker but longer lasting currents via metabotropic hcrt receptors.

The superposition of these two synaptic currents plays a major role for the excitation of thalamic neurons and especially for the activity state in the reciprocally connected neurons of the hypothalamus. The functionally most important component, thereby, is the activity-dependent change of the synaptic efficacy of hcrt transmission. This is considered by a modulation function M as an additional factor in scaling the parameter a of hcrt activation:

$$da_{hcrt}/dt = (a_{hcrt} M - a_{hcrt\infty})/\tau_{hcrt} \tag{7}$$

The synaptic efficacy, as represented by M, depends on the availability of presynaptic hcrt and/or postsynaptic receptors. It tends to achieve a maximum value $M_{\max}$ with time constant τ_{inc} while it is attenuated with each spike, by a certain amount which is scaled by the time constant τ_{dec}. Both decrease and recovery of the synaptic efficacy depend on its actual value M. This leads to the following equation:

Table 1 The parameters values for the model. For details see text and references [17, 18]

A: Voltage–gated currents

parameter / neuron	g_l	g_{Na}	g_K	$g_{Na,p}$	$g_{K,Ca}$	V_l	V_{Na}	V_K	$V_{0_{Na}}=V_{0_K}$	$V_{0_{Na,p}}$	$s_{Na}=s_K$	$s_{Na,p}$	τ_K
TH	0.1	1.3	1.75	0.22	0.35	−60	50	−90	−25	−40	0.25	0.09	3.5
HCRT	0.1	3.0	4.0	–	–	−60	50	−90	−25	–	0.25	–	2
GLU	0.1	3.0	4.0	–	–	−60	50	−90	−25	–	0.25	–	2

B: Postsynaptic currents

parameter / neuron	$\tau_{Na,p}$	$\tau_{K,Ca}$	η	k
TH	17	35	0.012	0.17

V_{syn}	V_{spike}	s_{syn}	τ_{glu}	τ_{hcrt}	τ_{dec}	τ_{inc}	M_{max}
50	−20	1	30	300	920	7500	1

$$dM/dt = (M_{\max} - M)/\tau_{dec} - a_{\infty}M/\tau_{dec} \tag{8}$$

The numerical values of the model parameters are given below. A detailed description of the equations of the hypothalamic model and their physiological meaning can be found in [14], and the details about the thalamic model in [15].

3 Simulation Results

In the above described model, the relevant mechanisms of sleep-wake transitions appear at the hypothalamic level and are related to the activity state of the HCRT neuron. According to the experimental data, firing of the HCRT neuron indicates a wake state and silence indicates sleep. The corresponding transitions and the associated changes of thalamic activity patterns are illustrated in Fig. 2.

The reciprocal feedback loop of the HCRT and the GLU neuron can be activated by an external stimulus which here is the circadian current input. Once the HCRT neuron fires strong enough to activate the GLU neuron, firing will be sustained due to the reciprocal excitation, also when the stimulus decreases or is completely eliminated.

The transition to a silent (sleep) state appears because of the decreasing synaptic efficacy (M) of hcrt transmission which leads to a decreasing firing rate. The firing rate is additionally decreased when the circadian input is simultaneously going down. Firing rate and synaptic efficacy, together, determine the activation variable a_{hcrt}. At the certain point, hcrt transmission is too weak to activate the GLU neuron and firing in the hypothalamic circuit stops. The activation a_{hcrt} drops to zero, while the synaptic efficacy starts to recover.

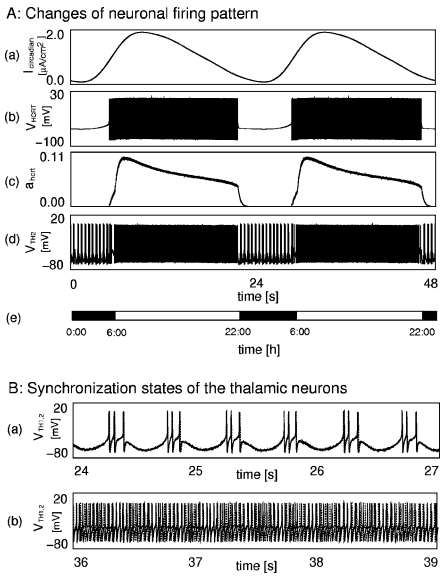


Fig. 2 Changes of the neuronal firing pattern in the hypothalamic and thalamic neurons according to sleep-wake cycles (**a**). Voltage potentials of the thalamic neurons, indicating their synchronization states (**b**). synchronized in bursting, asynchronous in tonic firing

When the circadian input reaches a certain value, firing can be reinstalled. Providing that the synaptic efficacy has sufficiently recovered, firing in both neurons can be sustained until it stops again due the decreasing synaptic efficacy as described above.

These transitions have significant influence on the activity patterns and synchronization states of the thalamic neurons. The excitatory synaptic input from the HCRT neuron keeps the thalamic neurons in a depolarized state with high frequent tonic firing activity, which is only slightly changed by the decreasing activation variable. When this input is missing, the thalamic neurons do not switch into a silent state. They are still active due to their intrinsic dynamics but the activity pattern changes to impulse groups (bursts). Alone these transitions from tonic firing to bursts are sufficient to bring the thalamic neurons from an unsynchronized to a synchronized state (Fig. 2b). These transitions correspond to the experimentally observed changes at the transitions from wakefulness to sleep.

4 Discussion

We have presented a mechanism-based model with neurons and synapses that connects hypothalamic mechanisms of sleep regulation with thalamic alterations of activity pattern and synchronization states at the transitions between sleep and wakefulness. These changes of thalamic activity seem to be closely associated with the control of conscious perceptions. Our model offers a simple but physiologically justified explanation how these transitions can appear. We propose that they are

introduced by the wake-active hypocretin neurons in the hypothalamus that project to the thalamic neurons where hcrt induces excitatory effects [12]. That is sufficient to introduce the typical sleep-wake-transitions related activity changes in the thalamic neurons. Everything else is the result of their intrinsic dynamics. No change of connectivity is required.

Without external current input the thalamic neurons are in a bursting mode while a depolarizing current can tune them into a tonic firing regime. In the bursting mode, the gap-junction coupled neurons easily synchronize, whereas they preferably stay in a desynchronized state in the tonic firing mode. Hence, while synchronization and de-synchronization is a question of the intrinsic dynamics of thalamic neurons, the transitions are introduced from outside, i.e. by synaptic input from sleep-wake controlling neurons in the hypothalamus.

In summary, we have illustrated a concept of functional interdependencies between thalamic gating mechanism and sleep-wake control by means of a mechanism-based neuronal model. In this way, we also have provided an example how the macroscopic level of systemic functions can be related to the microscopic level of cellular mechanisms, namely via a mesoscopic level of neuronal firing and activity pattern.

Acknowledgments This work was supported by the European Union through the Network of Excellence BioSim contract No LSHB-CT-2004-005137.

References

1. Blumenfeld, H., McCormick, D.A.: Corticothalamic inputs control the pattern of activity generated in thalamocortical networks. J. Neurosci. **20**(13) (2000) 5153–5162.
2. Bal, T., Debay, D., Destexhe, A.: Cortical feedback controls the frequency and synchrony of oscillations in the visual thalamus. J. Neurosci. **20**(19) (2000) 7478–7488.
3. Glenn, L.L., Steriade, M.: Discharge rate and excitability of cortically projecting intralaminar thalamic neurons during waking and sleep states. J. Neurosci. **2**(10) (1982) 1387–1404.
4. McCormick, D.A., Feeser, H.R.: Functional implications of burst firing and single spike activity in lateral geniculate relay neurons. Neuroscience **39**(1) (1990) 103–113.
5. Bazhenov, M., Timofeev, I., Steriade, M., Sejnowski, T.J.: Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. J. Neurosci. **22**(19) (2002) 8691–8704.
6. Hill, S., Tononi, G.: Modeling sleep and wakefulness in the thalamocortical system. J. Neurophysiol. **93**(3) (2005) 1671–1698.
7. Robinson, P.A., Rennie, C.J., Wright, J.J., Bahramali, H., Gordon, E., Rowe, D.L.: Prediction of electroencephalographic spectra from neurophysiology. Phys. Rev. E. **63**(2pt.1) (2001) 021903–021921.
8. Saper, C.B., Scammell, T.E., Lu, J.: Hypothalamic regulation of sleep and circadian rhythms. Nature **437**(27) (2005) 1257–1263.
9. Borbély, A.A.: A two process model of sleep regulation. Hum. Neurobiol. **1** (1982) 195–204.
10. Moore, R.Y.: Suprachiasmatic nucleus in sleep-wake regulation. Sleep Med. **3** (2007) 27–33.
11. Szymusiak, R., McGinty, D.: Hypothalamic regulation of sleep and arousal. Ann. NY Acad. Sci. **1129** (2008) 275–286.
12. Peyron, C., Tighe, D.K., van den Pol, A.N., et al.: Neurons containing hypocretin (orexin) project to multiple neuronal systems. J. Neurosci. **18**(23) (1998) 9996–10015.

13. Sakurai, T., Amemiya, A., Ishii, M., et al.: Orexins and orexin receptors: a family of hypothalamic neuropeptides and G-protein coupled receptors that regulate feeding behavior. Cell **92** (1998) 573–585.
14. de Lecea, L., Kilduff, T.S., Peyron, C., et al.: The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. Proc. Natl. Acad. Sci. USA **95** (1998) 322–327.
15. Lin, L., Faraco, J., Li, R., Kadotani, H., et al.: The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. Cell **98**(3) (1999) 365–376.
16. Chemelli, R.M., Willie, J.T., Sinton, C.M., et al.: Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. Cell **98**(4) (1999) 437–451.
17. Postnova, S., Voigt, K., Braun, H.A.: A mathematical model of homeostatic regulation of sleep-wake cycles by hypocretin/orexin. J. Biol. Rhythms **24**(6) (2009) 523–535.
18. Postnova, S., Voigt, K., Braun, H.A.: Neural synchronization at tonic-to-bursting transitions. J. Biol. Phys. **33**(2) (2007) 129–143.

Stochastic Resonance and Stochastic Encoding: Cooperative Effects of Noise and Intrinsic Dynamics in a Model Neuron with Subthreshold Oscillations

Wuyin Jin, Svetlana Postnova, and Hans Albert Braun

Abstract It is well established that noise, despite its mostly detrimental effects, can play a constructive role for information transfer. The best known example is stochastic resonance (SR) phenomena. SR means that there is an optimal level of noise for the detection of weak signals. In this study, a Hodgkin-Huxley type model neuron with subthreshold oscillations has been used to demonstrate that the noise effects essentially depend on the neuron′s activity state. These data shall emphasize more general aspects of stochastic encoding (SE) which, beyond tuning of noise, can be achieved with appropriate adjustment of physiological control parameters.

Keywords Nonlinear dynamics · Spiking neuron · Information transmission · Neuromodulation · Cross correlation · Subthreshold/suprathreshold stochastic resonance

1 Introduction

Neuronal information transmission is generally contaminated by noise which can come from numerous external and internal sources [1]. Although noise is usually considered as a disturbance of precise information transfer there are physiological situations where noise can play a constructive role. One of the best known examples is stochastic resonance (SR) that is an enhancement of signal encoding by an optimal level of noise (for review see [2].) SR effects have been observed in diverse physical and biological systems and also have been used to improve technical sensors [3, 4].

W. Jin (✉)
School of Mechanical & Electronic Engineering, Lanzhou University of Technology, Lanzhou 730050, P.R. China; Neurodynamics Group, Institute of Physiology, Philipps University of Marburg, Marburg D-35037, Germany
e-mail: wuyinjin@hotmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_86,

However, while technical sensors, as well as computer models, can artificially be tuned to optimal noise intensity, noise is not necessarily among the major neuronal control parameters which rather are ionic currents and conductances.

To examine the interdependencies between noise and neuronal dynamics for stimulus encoding we have used a Hodgkin-Huxley type model neuron with subthreshold oscillations [5]. Although the occurrence of SR in such model neurons has been denied [6], it has theoretically [7] and also experimentally [8] been demonstrated that noise, specifically in combination with subthreshold oscillations, can play a crucial role for stimulus encoding.

2 Methods: Model Equations

The simulations were made with a Hodgkin-Huxley type neuron which was originally developed for the examination of "stimulus sensitivity and neuromodulatory properties of noisy neuronal oscillators" [5]. The voltage changes dV/dt across the membrane capacitance C are calculated as the sum of specific ion currents:

$$C_M dV/dt = \sum I = -I_l - I_{Na} - I_K - I_{Na,p} - I_{Ks} - I_{ext} - I_{noise} \tag{1}$$

Apart from the leak current $I_l = g_l(V - V_l)$, with leak conductance g_l and the driving force $(V - V_l)$, there is a term for external current injection (I_{ext}) and for noise application (I_{noise}). The physiologically relevant dynamics are introduced by voltage-dependent currents. Two of them, I_{Na} and I_K, determine the shape of the action potentials. Two others, $I_{Na,p}$ and I_{Ks}, are for slow subthreshold oscillations.

All voltage-dependent currents are modelled in the same form with

$$I_i = g_i a_i (V - V_i), \quad i = Na; K; Na,p; Ks \tag{2}$$

The currents I_i are calculated as the product of the driving force $V - V_i$ (V_i is the equilibrium potential) and the maximum conductance g_i multiplied by a voltage- and time-dependent activation variable a_i. For simplicity, inactivation is neglected.

Time-dependent activation is modelled in form of first order differential equations

$$da_i/dt = (a_{i\infty} - a_i)/\tau_i \tag{3}$$

with time constants τ_i and steady state activation values $a_{i\infty}$ of sigmoid voltage dependencies, determined by the slopes s_i and half activation potentials V_{0i}:

$$a_{i\infty} = 1/(1 + \exp(-s_i(V - V_{0i}))) \tag{4}$$

3 Results

The model neuron has been tuned from constant membrane potential to subthreshold oscillations and to regular spiking by external current injection (Fig. 1a) The physiologically most relevant patterns, a mixture of subthreshold and spike generating oscillations, however, only develops with contribution of noise, thereby smoothing the deterministically step-like transfer function (Fig. 1b).

For systematic examination of stochastic resonance effects we have calculated the cross correlation coefficients (CCC) between a small sinusoidal current stimulus ($I_{ext} = 0.1sin(t/30)$, with I in $\mu A/cm^2$ and time in seconds) and the thereby induced alterations of the firing rate as a function of noise intensity (Fig. 2). For these simulations the neuron has been adjusted by constant base-line currents to different subthreshold (Fig. 2a) and suprathreshold (Fig. 2b) activity states. In this way, a series of typical SR curves with optimal noise levels could be obtained – with two exceptions. In each diagram there is one curve with highest CCC values for deterministic simulations. This is the case when the neuron the subthreshold stimulus can

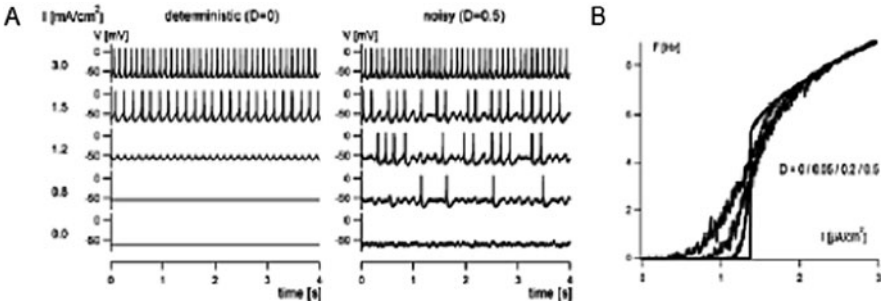


Fig. 1 Deterministic and noisy simulations in response to constant current injection illustrated by voltage traces (**a**) and frequency-stimulus curves (**b**)

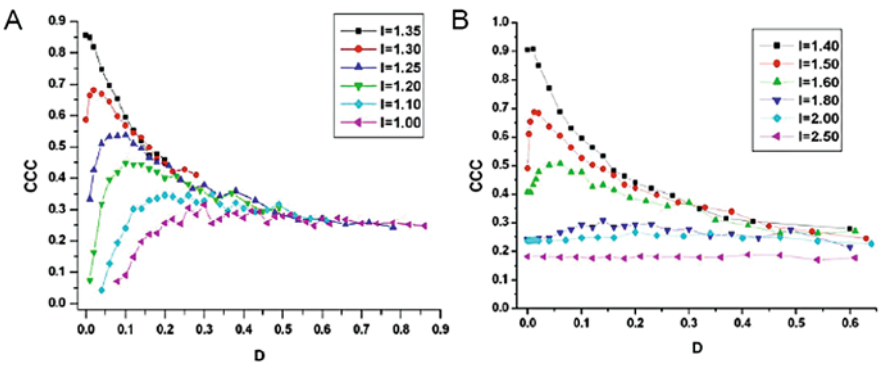


Fig. 2 Cross-correlation coefficients (CCC) between a sinusoidal current stimulus and the neuronal firing rate as a function of noise intensity D under deterministically subthreshold (**a**) and suprathreshold (**b**) conditions, adjusted by constant currents I

cross the deterministic threshold of spike generation without the help of noise. At lower base-line currents, noise is ultimately needed for stimulus encoding because otherwise the neurons would remain silent. At higher base-line currents, slight alteration of a deterministically ongoing firing rate can be enhanced by noise as long as it allows a modulation of the spiking probability

In comparison of the SR curves, it can clearly be recognized that the noise effects essentially depend on the base-line current: (1) the less noise intensity is required for optimal stimulus-response correlations the closer the neuron operates to the deterministic threshold for spike generation; (2) the efficacy of noise, as indicated by the peaks in the cross-correlation curves, is becoming less pronounced the further away the neurons operate from the deterministic threshold.

Nevertheless, at all subthreshold base-line current (Fig. 2a), noise of sufficiently high intensity still can improve stimulus encoding simply because the neuron otherwise would be completely silent. However, the SR effects, i.e. the maxima in the CCC curve, are more and more diminished. With suprathreshold currents (Fig. 2b), noise finally will become irrelevant as indicated by the completely flat CCC curve with I=2.50.

4 Discussion

We have illustrated stochastic resonance (SR) in a model neuron with subthreshold oscillations, for both sub- and suprathreshold conditions, and have demonstrated that optimal noise intensity essentially depends on the neuron′s operating point with regard to the deterministic threshold of spike generation.

Subthreshold oscillations can endow the neurons with particular tuning properties [7] which, in combination with noise, can lead to exquisite stimulus sensitivity [8]. The highest sensitivity is achieved at around 50% spiking probability of the subthreshold oscillations. The maximum sensitivity is reduced with increasing noise while the range of stochastic encoding (SE) is broadened. Likewise, the SR effects are becoming smaller the higher noise intensities are required for optimal stimulus-response correlation. Accordingly, it is the adjustment of the neuronal dynamics which decides about the neurons′ stochastic encoding properties including the eventual appearance of stochastic resonance phenomena.

Acknowledgments The study was supported by the NSFC of China, No.10572056, 30670529 and the EU through the Network of Excellence BioSim, No. LSHB-CT-2004-005137.

References

1. Faisal, A.A., Selen, L.P., Wolpert, D.M.: Noise in nervous system. Nat. Rev. Neurosci. **9**(4) (2008) 292–303
2. Gammaitoni, L., Hänggi, P., Jung, P., Marchesoni, F.: Stochastic resonance. Rev. Mod. Phys. **70** (1998) 223–288
3. Wiesenfeld, K., Moss, F.: Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. Nature **373** (1995) 33–36

4. Bulsara, A.R., Gammaitoni, L.: Tuning in to noise. Phys. Today **49**(3) (1996) 39–45
5. Huber, M.T., Krieg, J.C., Dewald, M., Voigt, K., Braun, H.A.: Stimulus sensitivity and neuromodulatory properties of noisy intrinsic neuronal oscillators. Biosystems **48** (1998) 95–104
6. Gong, Y., Xu, J., Hu, S.: Stochastic resonance: when does it *not* occur in neuronal models? Phys. Lett. A **243** (1998) 351–359
7. Huber, M.T., Braun, H.A.: Stimulus – response curves of a neuronal model for noisy subthreshold oscillations and related spike generation. Phys. Rev. E **73** (2006) 1–10
8. Braun, H.A., Wissing, H., Schäfer, K., Hirsch, M.C.: Oscillation and noise determine signal transduction in shark multimodal sensory cells. Nature **367** (1994) 270–273

Attention Modulation of Sensory Systems

Hans Liljenström and Yuqiao Gu

Abstract Attention and arousal seems to enhance the efficiency of neural information processing. We use computer simulations of various sensory systems in order to investigate how the neurodynamics of these systems can be modulated for optimal performance in an unknown and changing environment. Using an inter-scale model of the insect antennal lobe, we demonstrate how attention modulation can change the sensitivity for sex pheromones. We also study how neural oscillations in mammalian olfactory cortex can serve to amplify weak signals and sustain an input pattern for more accurate information processing, and how chaotic-like behaviour could increase the sensitivity in initial, exploratory states. Finally, we use a neural network model of visual cortex area V4, in order to investigate potential cellular and network mechanisms for visual attention, reproducing experimental findings of attention induced gamma-frequency synchronization.

Keywords Attention · Neuromodulation · Acetylcholine · Olfaction · Vision · Neural network models · Inter-scale interactions

1 Introduction

Sensory systems need to adapt to various environmental conditions and internal states. Attention and arousal seems to enhance the efficiency of the neural information processing in sensory systems of all kinds of animals, invertebrates as well as vertebrates. This efficiency enhancement is mediated by modulation of the neural activity at the cellular, network, and systems levels, thus involving interactions between micro-, meso-, and macro-scales. Important neuromodulators associated with attention are acetylcholine (ACh) and serotonin (5-HT), which may affect

H. Liljenström (✉)
Division of Biometry & Systems Analysis, SLU, Uppsala 750 07, Sweden; Agora for Biosystems, Sigtuna SE-193 22, Sweden
e-mail: hans.liljenstrom@et.slu.se

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_87,

the excitability of a large number of neurons simultaneously and the synaptic transmission between them [1, 2].

At least in mammals, attention plays a key role in perception, action selection, object recognition and memory. It is associated with primarily gamma frequency oscillations in cortical structures, for example in the visual cortex during visual attention [3, 4]. Data suggest that gamma rhythms are associated with relatively local computations, whereas beta rhythms are associated with higher level interactions. Generally, it is believed that lower frequency bands are generated by global circuits, while higher frequency bands are derived from local connections [4].

We study attention modulation of the neural activity in different sensory systems and use different types of models, depending on available data and the specific problem addressed.

2 Models

In order to study attention modulation in sensory systems, primarily for olfaction and vision, we use various types of neural network models: (1) of the insect antennal lobe, (2) of mammalian olfactory cortex, and (3) of visual cortex, area V4.

2.1 Neural Network of Insect Antennal Lobe

Natural odours are multidimensional and noisy signals, which in insects are processed in the antennal lobe (AL), the primary odour information processing centre that is organized into several glomeruli. All synaptic contacts between the three main types of AL neurons (olfactory receptor neurons, ORN, local inhibitory neurons, LN, and projection neurons, PN) take place within the glomeruli. The membrane potentials of PNs and LNs are described by the following dynamical equations, based on the standard Hodgkin-Huxley equations, with variables and parameters given in [5, 6]:

$$C_m \frac{dV_{PN}}{dt} = -g_L(V_{PN}-E_L)-I_{Na}-I_K-I_A-g_{KL}(V_{PN}-E_{KL})-I_{GABA_A}-I_{slow-inh}-I_{stim} \quad (1)$$

$$C_m \frac{dV_{LN}}{dt} = -g_L(V_{LN}-E_L)-I_{Ca}-I_{K(Ca)}-I_K-g_{KL}(V_{PN}-E_{KL})-I_{GABA_A}-I_{nACh}-I_{stim} \quad (2)$$

These equations take into account the ORN input (excitatory I_{stim}), the LN input (fast I_{GABA} , and slow $I_{\text{slow - inh}}$ inhibition), the PN input to LN (I_{nACh}) and the various other non-synaptic currents. Each model glomerulus has six PNs and two LNs.

2.2 Population Model of Olfactory Cortex

Our model of the olfactory cortex has network units with a continuous input-output function, corresponding to the average firing frequency of neural populations, which we compare with EEG and LFP data [7]. There are three cortical layers, with network units corresponding to populations of *feedforward* inhibitory interneurons, pyramidal cells, and *feedback* inhibitory interneurons, respectively. All connections are modeled with distance dependant time delays.

The time evolution for a network of N neural units is given by a set of coupled nonlinear first order differential delay equations for all the N internal states, u. With odour signal, I(t), noise ξ(t), characteristic time constant, τi, and connection weight wij between units i and j, separated by a time delay δij, we have for each unit activity:

$$\frac{du_i}{dt} = -\frac{u_i}{\tau_i} + \sum_{\mathrm{j}\neq\mathrm{i}}^{\mathrm{N}} w_{ij} g_j[u_j(t-\delta_{ij})] + I_i(t) + \xi(t) \tag{3}$$

The input-output function, $g_i(u_i)$, is a continuous sigmoid function:

$$g_i = C \cdot Q_i \left\{ 1 - exp\left[\frac{exp(u_i) - 1}{Q_i} \right] \right\} \tag{4}$$

The gain parameter Q_i determines the slope, threshold and amplitude of the curve for unit i. This gain parameter is associated with the level of arousal/attention as expressed through the level of acetylcholine. C is a normalization constant [7].

2.3 Network Model of Visual Cortex

In our model of visual cortex, we use spiking model neurons, since we want to compare our results with observed data, as spike triggered averages of local field potentials. All model neurons satisfy the following Hodgkin-Huxley equation:

$$\begin{aligned} CV' = &-g_L(V+67) - g_{Na} m^3 h(V-50) - g_K n^4 (V+100) \\ &-g_{AHP} w(V+100) - I^{syn} + I^{appl} \end{aligned} \tag{5}$$

where V is the membrane potential and C is the membrane capacitance. g_L is the leak conductance, g_{Na} and g_K are the maximal sodium and potassium conductances, respectively. g_{AHP} is the maximal slow potassium conductance of the afterhyperpolarization (AHP) current, which varies, depending on the attentional state: in an idle state, $g_{AHP} = 1.0$ ms, with attention, $g_{AHP} \leq 1.0$ ms. I^{syn} is the synaptic input current, and I^{appl} is the applied current. The variables m, h, n and w are calculated in a conventional way, and described more thoroughly in [8].

In each of the three (lumped) layers of the local area network, there are four types of interactions: (1) lateral excitatory–excitatory, (2) excitatory–inhibitory, (3) inhibitory–excitatory, and (4) inhibitory–inhibitory, with corresponding connection strengths, which vary with distance between neurons.

3 Simulation Results

3.1 Attention Modulation Enhances System Sensitivity

We connect four model glomeruli in a circle, to represent the male-specific sex pheromone processing system of the antennal lobe – the macroglomerular complex (MGC). In each simulation, we use a pulse as the pheromone input signal to stimulate one or two of the four glomeruli, depending on the problem addressed. The input to the four glomeruli in MGC during stimulation is represented by a four dimensional input vector, (P_1, P_2, P_3, P_4).

At the micro-scale, the PNs employ different ionic channels for coding. One of the features of the micro-scale circuit is that the potassium channels can be modulated by neuromodulators. We regard 5-HT as an *attention modulation signal* delivered from higher processing levels through a feedback pathway to the PNs in glomerulus 1, G_1, which processes the major pheromone component.

Our simulation results show that the threshold in response to a weak stimulus decreases and the excitability increases, when the maximal conductance of the I_A of the PNs in G_1 is decreased. This suggests that *attention increases the sensitivity* of PNs in G_1. We also find that attention to the major pheromone component can suppress the stimuli to the other glomeruli, due to the local attention modulation and lateral global inhibition of the network. Figure 1 shows that the system can increase the response excitability to the major pheromone stimulus, while decreasing the response excitability to stimulus input with the same dose to G_2.

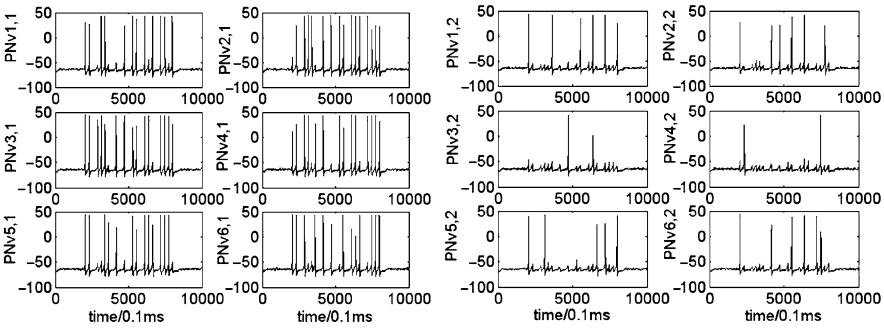


Fig. 1 Attention modulation of G_1 increases the response excitability to the major pheromone component stimulus (*left*), and decreases the response excitability to the stimulus of G_2 (*right*). The stimulus vector is here (300, 300, 0, 0)

3.2 Neuromodulation Enhances Learning and Memory

In the case of mammalian olfactory cortex, we simulate neuromodulatory control by varying the gain parameter Q, and by varying excitatory and inhibitory synaptic transmission (reducing w_{ex} and w_{in}). In the first case, neuronal excitability is affected, corresponding to the level of arousal or attention of the animal, chemically linked to a neuromodulator. The second effect is characteristic of ACh [2].

Different values of these parameters can result in different dynamical states: static, oscillatory or chaotic. For example, by changing the parameter Q (in Eqn. (4)) from 1 to 15, the dynamics changes from static to oscillatory to chaotic states. At an intermediate value of Q, the system may have a long transient chaotic phase, which converges to an oscillatory state, corresponding to a limit cycle memory state. This behaviour seems advantageous in allowing for an initial sensitivity to input signals, a quick search of state space, until an eventual convergence to one of the stored memories most closely resembling the input data. In Fig. 2, such a process is illustrated by the trajectories in state space, following the time evolution of two arbitrary network units, u_1 and u_2, plotted against each other. Note the initial irregularity, which slowly converges into regular cycles [9].

The frequencies of the network oscillations depend primarily upon intrinsic time constants and delays, whereas the amplitudes depend predominantly upon connection weights and gains, which are under neuromodulatory control. Implementation of these neuromodulatory effects in the model caused changes analogous to those seen in physiological experiments [10]. In particular, the increase in excitability and the suppression of synaptic transmission can induce theta and/or gamma oscillations within the model, even when starting from an initially quiescent state. The increase in excitability could also result in an enhancement of gamma frequency oscillations in the response to an external input.

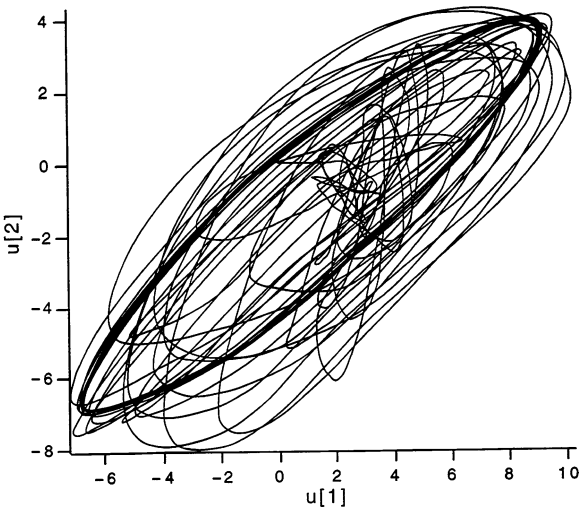


Fig. 2 Transient chaos obtained for $Q = 14.0$. The activity of two excitatory units is plotted against each other for 1 simulated second. After the initial chaotic-like phase (approximately 500 ms), the activity converges to a near limit cycle state, corresponding to an approximately 40 Hz oscillation

3.3 Attention Effects During Visual Stimulation

To mimic the situation in a visual attention experiment [4], we have groups of "attended-in" neurons, A_{in} (where attention is directed to a stimulus location inside the receptive field (RF) of these neurons) and groups of "attended-out" neurons, A_{out} (where attention is directed to a stimulus location outside the RF these neurons). During a stimulus period, two identical stimuli are presented; one appears at a location inside the RF of A_{in} and the other appears at a location inside the RF of A_{out}. To simulate the dynamics during a stimulus period, we apply bottom-up sensory stimulation currents, as well as top-down attention (cholinergic) modulation.

Figure 3 shows the simulated spikes of one pyramidal neuron in A_{in} and of one pyramidal neuron in A_{out}, as well as the LFP, STA and STA power of the A_{in} and A_{out} neurons in layer 2/3. In comparison with an idle state, the dominant frequency of the STA power spectrum of both A_{in} and A_{out} neurons shifts to gamma band, due to the stimulation inputs. Also, gamma synchronization in A_{in} is always stronger than in A_{out} neurons. This result agrees with the experimental findings that gamma synchronization increases in A_{in} during a stimulus period [4].

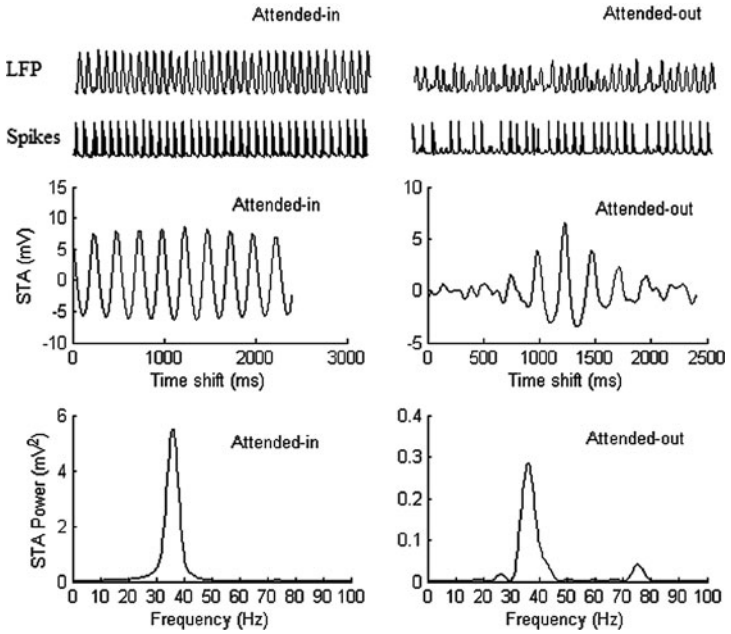


Fig. 3 Attention modulation effects during a stimulus period. LFP, spikes, STA and STA power of *attended-in* and *attended-out* group calculated for the superficial layer, when the excitatory connections and inhibitory connections to each pyramidal neuron in the attended-in group within R_{modu} in layer 2/3 and layer 5/6 are reduced to zero. Note the different scales in the y axes, and see text for further details

4 Discussion

We have used different types of neural network models of sensory systems to investigate how attention, by means of various neuromodulators, may affect the efficiency of neural information processing. Our model simulations show that many factors play important roles in the network neurodynamics. These include (1) the interplay of ion channel dynamics and neuromodulation at a *micro*-scale, (2) the local network connectivity pattern, (3) the feedforward and feedback connections between different network structures at a *meso*-scale, and (4) the top-down and bottom-up circuitries at a *macro*-scale.

In conclusion, the dynamical state of a neural system determines its global properties and functions. Neurodynamic control, or modulation should thus be crucial for the survival of an animal, or for an efficient functioning of any autonomous system. It could result in a shift in the balance between sensitivity and stability of the system. While the functional role and control of the neurodynamics is still largely an open question, computer models can provide a useful tool to address these problems.

References

1. Freeman, W.J.: Neurodynamics: An Exploration in Mesoscopic Brain Dynamics. Berlin: Springer (2000)
2. Liljenström, H., Hasselmo, M.: Cholinergic modulation of cortical oscillatory dynamics. J. Neurophysiol. **74** (1995) 288–297
3. Gray, C.M., Singer, W.: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc. Natl. Acad. Sci. USA **86** (1989) 1698–1702
4. Fries, P., Reynolds, J.H., Rorie, A.E., Desimone, R.: Modulation of oscillatory neuronal synchronization by selective visual attention. Science **291** (2001) 1560–1563
5. Bazhenov, M., Stopfer, M., Rabinovich, M., Huerta, R., Abarbanel, H.D.I., Sejnowski, T.J., Laurent, G.: Model of transient oscillatory synchronization in the locust antennal lobe. Neuron **30** (2001) 553–567
6. Gu, Y., Liljenström, H.: Modeling efficiency in insect olfactory information processing. BioSystems **89** (2007) 236–243
7. Liljenström, H.: Modeling the dynamics of olfactory cortex using simplified network units and realistic architecture. Int. J. Neural Syst. **2** (1991) 1–15
8. Gu, Y., Liljenström, H.: A neural network model of attention-modulated neurodynamics. Cogn. Neurodyn. **1** (2007) 275–285
9. Wu, X., Liljenström, H.: Regulating the nonlinear dynamics of olfactory cortex. Network Comput. Neural Syst. **5** (1994) 47–60
10. Biedenbach, M.A.: Effects of anesthetics and cholinergic drugs on prepyriform electrical activity in cats. Exp. Neurol. **16** (1966) 464–479

Evidence for a Spatiotemporal Singularity in Percept Formation by Cerebral Cortex

Walter J. Freeman

Abstract Perception proceeds by cinematic frames that are separated by an endogenous *shutter*, which is intrinsic to the background *spontaneous* activity of cerebral cortex. Cinematic displays of spatiotemporal electrocorticograms (ECoG) show turbulence, from which stable spatial patterns emerge that are correlated with perception of conditioned stimuli (CS). Extracting invariant properties of the perceptual patterns requires high-resolution measurements by high-density electrode arrays, high-order FIR filters, and the Hilbert transform giving analytic signals (instantaneous amplitudes and frequencies). Each pattern is preceded by a *null spike*, which is postulated to reveal a *singularity* in cortical dynamics that operates the shutter. A piecewise linear model is used to solve the nonlinear differential equations of mesoscopic cortical dynamics. Solutions point to the existence of a limit cycle attractor. The behavioral and physiological circumstances are described by which the trajectory of cortical dynamics may converge to the attractor and precipitate a phase transition, which destroys an existing pattern and initiates a new one.

Keywords AM pattern · Electrocorticogram (ECoG) · Null spike · Perceptual frame · Phase transition · PM pattern (cone) · Singularity · Vortex

1 Introduction

Sensory cortices in brains sustain the embedding of the self in the environment by organizing sensory information into perceptions that express the meaning and significance of the sensory input. Each batch of sensory input from a saccade, whisk or sniff of a conditioned stimulus (CS) delivers a volley of action potentials to the appropriate sensory cortex. That burst of input helps to trigger a phase transition [1, 2] that condenses a disorderly receiving phase into a widely synchronized,

W.J. Freeman (✉)
Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
e-mail: dfreeman@berkeley.edu

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_88,

orderly transmitting phase. The order is expressed in a narrow band carrier wave that is modulated in amplitude (AM) and phase (PM) [3–5]. The AM and PM patterns are fixed within *frames* between phase transitions [6]. The spatial AM pattern is governed by an attractor that is selected by the CS from a landscape of basins of attraction. Its detailed shape is determined by the synaptic connections among cortical neurons in a nerve cell assembly that stores the memory of the CS as a category. That recalled memory is transmitted by the sensory cortex to other parts of the brain in the form of the AM pattern. The sensory information updates the synaptic connections in the assembly and is then discarded. The spatial PM pattern in the form of a phase cone defines the spatial location and temporal duration of the synchrony.

The hypothesis is proposed [7, 1] that the initiation of a phase transition requires the approach of cortical dynamics to a singularity [8] that is manifested by an extreme event [7, 9, 10] at a *point in time and space*. This extreme event called a *null spike* [7] has been identified as a sudden decrease in the background activity that occupies the spectral band of the induced carrier wave. New data confirm the extreme space-time localization of null spikes that precede the onsets of AM and PM patterns (phase cones). The spatial location of null spike and the apex of the following cone often coincide, but only after optimization of the search parameters and avoidance of overlapping wave bursts. The singularity is explained as a limit cycle attractor [11], which the cortical dynamics approaches in the conjunction of several conditions: high behavioral arousal, an increase in amplitude of cortical activity from a CS, and the approach toward zero in the power of the background filtered ECoG: the null spike. The conjunction increases the signal-to-noise ratio without limit. By the crossing of an as yet ill-defined threshold, a disorderly "gaseous" receiving phase condenses into an ordered "liquid" transmitting phase by a phase transition [12].

2 Methods

2.1 Experimental Methods

ECoG data were recorded from the olfactory bulb and cortex and the visual, auditory and somatic cortices of rabbits that were trained to respond to olfactory, visual, auditory or somatic CS [13]. High resolution was required for measurement of ECoG signals in the spatial, spectral, and temporal dimensions. High *spatial resolution* was optimized by use of spatial spectral analysis to minimize the distance between electrodes (0.79 mm) in square electrode arrays with optimal aperture (6×6 mm) on the cortical surface (Chapter 4 in [11, 14]). High *spectral resolution* was by a FIR filter with order 1,000–4,000 to optimize a pass band of 5 Hz in the beta and gamma ranges [7, 15]. High *temporal resolution* was optimized by sampling at 500/s over time segments lasting 6 s, which was the duration of individual trials (3 s control period with no CS followed by 3 s with CS) and by use of the Hilbert transform [16]. The *instantaneous analytic signal* was calculated and decomposed into the analytic

Table 1 Filter settings based on Nyquist frequency (Nf)

Band	Band pass (Hz)	2 ms step	5 ms step	Display, s
High gamma	50–100	1/2.5–1.5 Nf	1/2–1 Nf	1
Low gamma	25–50	1/10–1/5 Nf	1/4–1/2 Nf	2
Beta	12.5–25	1/20–1/10 Nf	1/8–1/4 Nf	4
Alpha	6.25–12.5	1/40–1/20 Nf	1/16–1/8 Nf	8
Theta	3.125–6.25	1/80–1/40 Nf	1/32–1/16 Nf	16

power (amplitude squared) and the unwrapped analytic phase [3, 4]. The analytic frequency was calculated as a time series by dividing successive analytic phase difference (in radians) by the duration of the digitizing step (here 0.002 s) and by 2π radians/cycle to express the values in Hz.

The power spectral density was calculated with the multitaper window [17] (MATLAB pmtm) and displayed in log-log coordinates. The choice of the order of the filter depended on the digitizing step of 2 ms and the pass band. Empirically order 500 gave invariance of the distribution of minimal amplitudes in null spikes over the relevant pass bands (Table 1) [15]. The pass bands were calculated in multiples of the Nyquist frequency (100 and 250 Hz) in order to approximate the conventional pass bands, and in order to show the scale-invariance of the oscillations in each pass band, when the duration of display was changed inversely with the frequency range [15].

2.2 Theoretical Methods

The simplest mesoscopic cortical circuit that suffices to model cortical dynamics contains two neural populations that form a KII set [18]: one is an excitatory KIe set; the other is an inhibitory KIi set. The minimal connections require that each KI set be modeled by a positive feedback loop between respectively 2 KOe sets and two KOi sets. The linear part of the KO cortical dynamics is revealed by the non-oscillatory *impulse response* (averaged evoked potential) under deep anesthesia. It can be fitted with the solution of a 2nd order ordinary differential equation (ODE), which is the sum of two exponentials. One term gives the initial rise rate, $a = 220/s$, and the other term gives the decay rate, $b = 720/s$, to the pre-stimulus baseline [11].

In the waking, behaving state the impulse response is oscillatory; it is measured by fitting to it the sum of two damped cosines [11, 18]. The power-law distribution of connection density with distance of transmission [19] endows the cortex with the property that, when a phase transition begins, it extends from a site of nucleation over the entire sensory cortex irrespective of its size. The resulting long correlation length makes it feasible to model the transition dynamics with a lumped approximation of the 2-D system. The linear part of the KII set is modeled by four 2nd order ODE [18]. The two rate constants, a and b, are the same for all four KO sets. The nonlinear part of the KII set is modeled by a time-invariant asymmetric sigmoid function at the output of each of the four populations.

The system is piece-wise linearized by replacing the sigmoid curve of each KO set with 1 of 4 fixed gain coefficients that are equal to the slope of the tangent to the sigmoid curve at a given operating point. That point is determined by measurement of the closed loop impulse response. It has the form of the sum of 2 damped cosines. The four interactions of the KII set are expressed in four nonlinear gain parameters: excitation of excitatory or inhibitory, inhibition of excitatory or inhibitory.

The important parameters from fitting damped cosines to the impulse responses are the amplitude, frequency, ω, and decay rate, α, of the cosine with ω in the gamma range and the value of α that is closest to the $j\omega$ axis. The 4 gain coefficients in the model predict the values of ω, α, amplitude, v, and phase, ϕ, of the impulse response. Conversely, the experimental values of ω, α, v, and ϕ serve to evaluate the 4 gain coefficients in the model.

Sets of solutions are generated to calculate the changes in the impulse responses with changes in the state of the cortex that are expressed in the closed loop poles and zeroes of the linear solutions. The state dependence of cortical dynamics is displayed by the paths of the closed loop poles and zeroes that are generated by changing the cortical state with anesthetics or changing the intensity of the input impulse, $\delta(t)$. These paths are *root loci*. Owing to the symmetry of the upper and lower halves of the complex plane, only the upper half need be displayed. Full details of methods are available on-line [11].

3 Results

3.1 Experimental Results

The temporal ECoG power spectral density (PSD_T) in coordinates of log power vs. log frequency was calculated over segments relatively long in duration (1–6 s). Typically the PSD_T conformed to a power-law (1/f) distribution in three frequency ranges having different slopes (Fig. 1A, *black curve*). In the range of delta, theta and alpha (<12.5 Hz) the slope tended toward zero ($1/f^0$) but with peaks above the trend line in segments having narrow band oscillations (Fig. 1A, *dark curve*). In the range of high beta and most of gamma (12–80 Hz) the slope varied between the limits of -1 to -3 ($1/f^1$ to $1/f^3$) depending on the behavioral state (especially waking vs. sleeping [20]). Above 80 Hz the slope was close to -4 ($1/f^4$) in the absence of noise [21] and to 0 ($1/f^0$) in the presence of noise [22].

The PSD_T of short segments (80–120 ms) extracted with a non-overlapping window stepped across the ECoG revealed multiple peaks coexisting at each step (Fig. 1A, *gray curve*), which varied in center frequency from each step to the next. The PSD_T was used to decompose the ECoG by centering a narrow band pass filter (5 Hz width) on each peak. The output was an oscillation that waxed and waned in amplitude in correspondence with the formation of bursts in the ECoG at different carrier frequencies. The example shown in Fig. 1C, was from the peak in power in the high gamma range (50–55 Hz). During high power (Fig. 1B) the spatial pattern

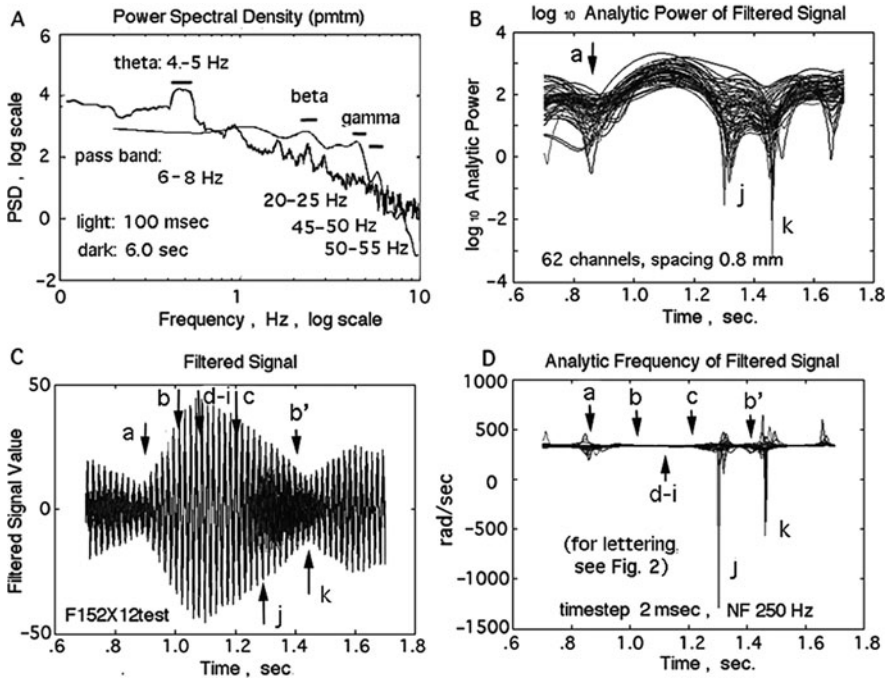


Fig. 1 **A** PSD_T. Finding the carrier band to within ± 2.5 Hz was crucial [20]. **B** Gamma ECoG, 62 superimposed signals. **C** Log_{10} analytic power from the Hilbert transform of band pass filtered signals [16]. **D** Analytic frequency. Phase resolution needed to demonstrate spatial coincidence of the null spike and the apex of the following conic phase gradient holds only for non-overlapping down spikes (as at **a**) and not when other bursts superimpose (as at **j** and **k**)

of amplitude and the frequency (Fig. 1D) both were relatively fixed [3] but they both changed from each frame to the next.

Evidence for the mechanism of the change was sought in the amplitude trough of between frames. The trough was visualized by superimposing all available signals (62 channels in Fig. 1). The display worked because all the signals had the same instantaneous frequency with minimal phase dispersion (within $\pm\pi$ rad $= \pm 45°$). At one or more time steps the analytic power (B, a) decreased to very near zero at or near one channel, and the analytic frequency (D, a) was indeterminate (as manifested by the very high spatiotemporal variance). At that moment there was a discontinuity in the ECoG signal in this pass band that was expressed in phase slip from one pre-existing carrier frequency in a preceding frame to a new carrier frequency in the next succeeding frame on all signals. This event (downward spike in poser, maximum spatial SD_X of the analytic frequency, and analytic phase discontinuity) was inferred to serve as a *temporal marker* for the phase transition.

A *spatial marker* for the phase transition was revealed by plotting the analytic power as a function of spatial location (Fig. 2). Serial spatial maps of the log_{10} analytic power in the filtered ECoG (the time series in Fig. 1C), when displayed as

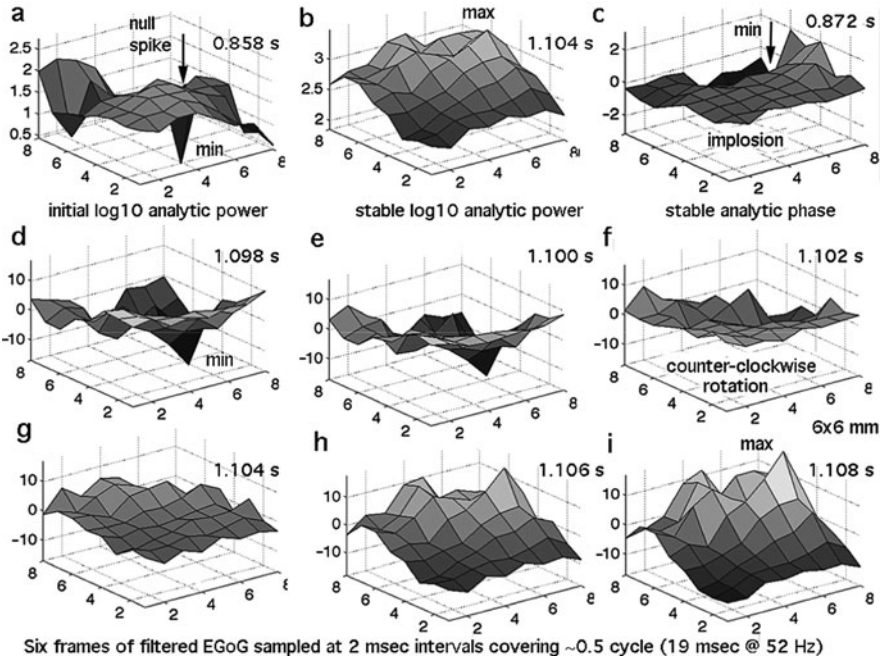


Fig. 2 Co-localization was found of the null spike, cone apex, and center of rotation in bursts in which one frequency dominated (Fig. 1, **C**). **a** The spatiotemporally localized null spike at **a** in Fig. 1 emerged between AM patterns. **b** The invariant AM pattern of the burst from **b–b′** in Fig. 1 was determined by learning. **c** The invariant PM pattern in **b–c** in Fig. 1 conformed to an inverted cone. **d–i** Half cycle ECoG amplitudes at 52 Hz revealed rotation [1, 2]

movies [23], revealed a surface that flapped and rolled like an ocean in storms [1, 2]. Intermittently the power decreased in a brief spike (Fig. 2A). The shape and width of the spike (Fig. 2A) conformed to the point spread function of a point dipole located at a depth of 1–1.5 mm below the recording array on the cortex, which corresponded to the depth of the generating layer of cortical dendrites (Chapter 4 in [11, 6]).

Another type of spatial marker had previously been found by mapping the 64 phase values of the carrier wave at its center frequency with respect to the spatial average phase at each time step [13, 14]. The phase surface formed by the 8×8 array of phase values was fitted with a cone. In roughly half of the cones the apex had phase lead corresponding to an explosive radiation, whereas half had phase lag at the apex (implosion). The location of the apex of a cone was postulated to correspond to the maximum or minimum of phase and to coincide with the location of the preceding null spike. It did so (Fig. 2C) if there was only one dominant frequency in the oscillatory burst (time segment **a** to **b′** in Fig. 1C) but not if the segment contained two components or more (preceding **a** and segment **j** to **k**). It was postulated that two or more superimposed bursts distorted the phase distribution from a cone [4]. Attempts to fit the sum of two cones failed, owing to insufficient information for convergence of nonlinear regression [14].

A third source of evidence for a critical spatial point in the phase transition came from the cinematic display [23] of the filtered ECoG amplitude (Frames d–i in Fig. 2). The spatial pattern of the ECoG amplitude in a single half-cycle was sampled from minimum to maximum at 2 ms time steps at the arrow d–i in Fig. 1C). The center frequency was 52 Hz; the wavelength was 19 ms, so each half cycle had six 2-ms steps. The successive cinematic frames [23] in this example revealed counterclockwise rotation of the field of ECoG potential around a center point close to the site of the preceding null spike. The field resembled the appearance of a vortex in a satellite image of a hurricane [24, 1]. Other examples revealed clockwise rotation, irrespective of the sign of the phase gradient (negative: explosion: positive: implosion). In contrast, the normalized spatial pattern of the analytic power (Fig. 2b) was constant across the half cycle and in fact across the time segment b to c in Fig. 1b. Collocation of the center of rotation with the null spike preceding the frame was not observed unless there was only one dominant frequency of oscillation carrying a spatial AM pattern.

3.2 Theoretical Results

The solutions to the piece-wise linear approximation of cortical dynamics by four 2nd order ODE generated 8 poles. Of the 8 poles the only pair of importance was the complex conjugate pole pair closest to the imaginary axis of the complex plane, and then only if the root loci crossed the imaginary jω axis. Owing to the symmetry of the complex pole pair, $-\alpha \pm j\omega$, the display of the root locus plot was restricted to the upper half of the complex plane in the frequency range of $\omega = 0$–400 rad/s (64 Hz) and decay rate $\alpha = -200$/s to +100/s. The graph showed a family of root loci. Each locus running mainly down and rightward showed the decreases in (ω, α) with increasing response amplitude (shown by the tick marks). The increase in amplitude imposed a concomitant increase in negative feedback gain, as predicted from the sigmoid nonlinearity (Chapter 3 in [11, 18]). The group of root loci (stacked from left to right) showed the increase in frequency and amplitude with increase in the steepness of the sigmoid curve with the degree of arousal (Chapter 5 in [11]), resulting in increased forward gain.

The conditional instability of the cortex indicating the possibility of a phase transition was predicted by the crossing of the group of root loci over the imaginary axis. The upper crossing was from left to right (conversion of the sign of the real part, α, of the root was from $-$ to +) with increasing forward gain and negative feedback gain. The crossing predicted that response amplitude would continue to increase and carry the system into an exponential increase in oscillatory amplitude: a burst. However, the root loci turned to the left and crossed back into the left of the jw axis. The convergence indicated an approach to a limit cycle attractor, because further increase in amplitude would increase the decay rate, a, and thereby decrease the feedback gain. This piecewise linear analysis provided a scaffold for using the impulse response to demonstrate the relation between degree of stability of the cortex modeled by the KII set in relation to arousal, sensitization by learning forming Hebbian assemblies for CS [5], and the predicted frequencies of bursts.

What remained to be explained sorted out was the sources of the spontaneous variation. The group of root loci could be explained by fluctuations in arousal, including the trend in each experimental data set for reduced arousal from satiety and fatigue. The orthogonal variation along each root locus in the group was replicated by giving sufficient anesthetic to reduce the background activity to a low level and then augmenting it by tetanic stimulation (high-frequency electrical stimulation, e.g., 200 Hz) of the incoming axons to the sensory cortex. This artificial noise temporarily reversed the effect of the anesthetic by increasing the frequency of oscillation and the decay rate over the physiological beta and gamma ranges (Fig. 4). Therefore the spontaneous variation was at least in part attributed to the fluctuations in the background Rayleigh noise in the filtered ECoG, including the down spikes. In brief, the location of the operating point of the cortex as revealed by the frequency and decay rate of the impulse response was dependent on arousal and on the signal-to-noise ratio of the CS-induced signal to the intrinsic, self-organized background spontaneous noise.

4 Discussion

The proposed hypothesis is that the phase transition that initiates each new frame in perception is dependent on a singularity in cortical dynamics. The key experimental evidence is the null spike, which has been discovered by pushing to the limits of the spatial, temporal and spectral resolution of ECoG. Experimentally the singularity resides in a spatiotemporal point, which does not conform to a pontifical neuron or to a cortical column but to a unique point in a vector field of mesoscopic activity [24]: the null spike; the conic apex; the center of rotation. The key theoretical evidence comes from the group of root loci fin the piecewise linear analysis of the KII set. The onset of instability is marked by the sign reversal (− to +) of the real part, a, of the complex pole pair, $-\alpha \pm j\omega$, with increasing amplitude. The singularity is modeled by the convergence of the root loci where they converge to the $j\omega$ axis and $\alpha = 0$. It occurs when the operating point approaches the limit cycle attractor, where the amplitude predicted from linear approximation increases exponentially, the background noise approaches 0, and the signal-to-noise ratio approaches ∞.

4.1 A Proposed Neural Mechanism of the Null Spike

The first question for discussion is how the null spike is generated. The answer is found in conceiving the background activity in the rest state as band-limited noise that generates beats: Rayleigh noise. The background ECoG has been shown to be close to brown noise, for which the PSD_T is power-law with a slope $= -2$ ($1/f^{\epsilon}$, $\epsilon = 2$). That would hold only for the output of an ideal random source that was free of other kinds of noise. The power-law PSD_T of ECoG was found to have slopes that were closer to -3 (*black noise*). That deviation from the ideal was shown to be due

to the intrinsic stabilization of the level of the background by the refractory periods [20] that decreased power in the PSD_T in proportion to the square of the frequency.

The form of the impulse response (the damped cosine) implied that in the awake cortex showed that the negative feedback interactions between the excitatory and inhibitory populations operated as a band pass filter, which revealed the characteristic frequencies in the beta and gamma ranges, because that perturbation, $\delta(t)$, contained all frequencies. The action of this intrinsic filter on the background activity in active states was seen in the oscillatory bursts and peaks in the PSD_T at high temporal resolution of the spectrum (Fig. 1, A). In the absence of input the endogenous oscillations were indistinguishable from the broad spectrum of black noise. The filtering action of the KII negative feedback that was simulated by applying band pass filters to the ECoG gave Rayleigh noise. The action of the negative feedback on the intrinsic oscillations evoked by sensory input was to give Rice noise [25], for which the intervals between beats were up to twice as long as those for Rayleigh noise. The null spikes occur as beats between the bursts [7].

By experiment [7] and theory [25] the modal interval between down spikes was a function solely of the bandwidth at all center frequencies. The modal interval that conformed to the repetition rate of bursts in the theta range (3–7 Hz) predicted the optimal pass band of 5 Hz. This value was confirmed experimentally. The bandwidth of the KII negative feedback was estimated from the minimal spatial standard deviation (SD) of the analytic frequency in each frame [7].

The distribution of the magnitudes of the down spikes in the rest state was random, because it conformed to that of spikes in filtered black noise [15]. Deviations from randomness in distribution from the active state appeared as excesses and deficits from values predicted for random noise [15, 9]. The modal interval between bursts that had classifiable AM patterns suggested that the threshold for reduction in power in the null spikes that would incur phase transitions was ~10^{-4} below the modal power [15].

4.2 Interpretation of the Role of the Null Spike in Perception

These diverse findings give a testable hypothesis of how the enormous neural population in a sensory cortex might transform itself from each frame to the next in an eye blink, destroying a pre-existing pattern, and initiating a new one that conveys a meaningful perception based on current sensory information. The key to interpretation is to consider the ECoG not as an electric potential but as an order parameter [26], that is, as an *index* of the strength of mesoscopic interactions among a cortical population. That interaction enables a sensory cortex to integrate its sensory information with its past experience and create an active memory that drives behavior. The ECoG order parameter is a vector that is evaluated by the AM pattern in each frame [24, 8].

The mass action that results from interaction must be sufficiently powerful to exert its influence on large populations elsewhere in the brain. This is ensured by the global synchrony of simultaneous transmission by virtually all of the projection

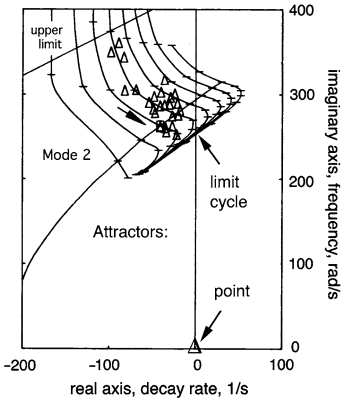


Fig. 3 The plot symbols Δ show the frequency, ω, and decay rate, α, of the damped cosine fitted to the impulse response of the olfactory cortex. The cluster of points came from a set of responses at fixed input intensity; the variation reflects the fluctuations in the level of the background activity. The curves represent a family of root loci, which show the direction of decrease in frequency and decay rate (*arrow downward to the right*) with increase in response amplitude. The two attractors are inferred to govern the cortical dynamics. From [11, p. 374, Fig. 6.21(**a**)]

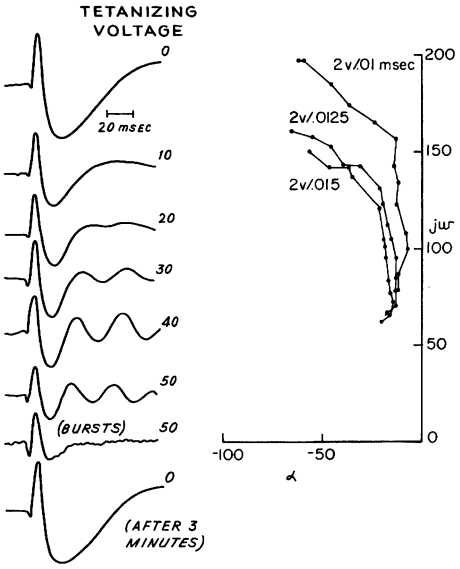


Fig. 4 On the left are examples of impulse responses (averaged evoked potentials) from the cat olfactory cortex. The background was suppressed by anesthesia and then replaced by tetanizing the axons. The frequency, w, and decay rate, a, increased in proportion to the intensity of the background activity, showing that the stability of cortex was reduced (*decrease in decay rate*) by the decrease in the background level, putatively resulting from the spontaneous occurrence of down spikes. From [11, p. 365, Fig. 6.13–14]

neurons in the cortex and by a spatial integral transformation that is performed by a divergent-convergent topology in the pathway carrying cortical output [19]. Further evidence for the repetitive framing of brain activity patterns has been found recently through extension of the search for synchronized amplitude and phase patterns from the intracranial ECoG to the scalp EEG [10, 27, 28].

Pursuit of this speculative hypothesis will require both experimental and theoretical advances. The main experimental deficiency is that the null spike usually fails to coincide with two other singular values: the apex of a conic phase distribution and the center of rotation of a vector field. The failure can be ascribed to several factors, including undersampling the spatial and temporal data streams, inadequate and inaccurate spatial and temporal decomposition of the AM and PM patterns, and lack of robust descriptions of null spikes, which often appear to move in their brief lifespan, which occur in clusters that are analogous to tornado vortices in weather systems [24, 1], and which are too brief to detect rotation at existing digitizing rates.

The main theoretical deficiency is the lack of mathematical structure linking the basic physics of singularities in vector fields with the electrophysiological data, with linear control systems in electrical engineering, and with the theory of behavior. Three behavioral factors determine the approach of the operating point along a root locus to the attractor that expresses the dynamic of phase transition. One is arousal contributing to the signal amplitude (Fig. 3) and the background noise level (Fig. 4). The second is the input-dependent amplification of the impulse response by the sigmoid nonlinearity. The third is the selective sensitivity to a CS that depends on formation of a Hebbian cell assembly through learning. The concept of singularity is proposed here to indicate a fruitful direction [29] in which to explore a major question in contemporary neuroscience: how do we perceive?

References

1. Freeman, W.J., Vitiello, G.: Dissipative neurodynamics in perception forms cortical patterns that are stabilized by vortices. J. Phys. Conf. Ser. **174** (2009) 012011.
2. Kozma, R., Freeman, W.J.: Intermittent spatio-temporal de-synchronization and sequenced synchrony in ECoG signals. Chaos 18 (2008) 037131.
3. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin. Neurophysiol. **115** (2004) 2077–2088.
4. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 2. Analytic phase. Clin. Neurophysiol. **115** (2004) 2089–2107.
5. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. Clin. Neurophysiol. **116/5** (2005) 1118–1129.
6. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 4. Analytic phase. Clin. Neurophysiol. **117/3** (2006) 572–589.
7. Freeman, W.J.: Deep analysis of perception through dynamic structures that emerge in cortical activity from self-regulated noise. Cogn. Neurodyn. **3**(1) (2009) 105–116.
8. Freeman, W.J., Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Phys. Life Rev. **3** (2006) 93–118.
9. Freyer F, Aquino K, Robinson PA, Ritter P, Breakspear M (2009) Bistability and non-gaussian fluctuations in spontaneous cortical activity. J. Neurosci. **29**(26): 8512–8524.

10. Pockett, S., Bold, G.E.J., Freeman, W.J.: EEG synchrony during a perceptual-cognitive task: Widespread phase synchrony at all frequencies. Clin. Neurophysiol. **120** (2009) 695–708 doi:10.1016/j.clinph.2008.12.044.
11. Freeman, W.J.: Mass Action in the Nervous System. New York, NY: Academic Press. For the singularity, see equations in Chapter 6 and summary Fig. 6.30 on p. 388 (1975/2004) http://sulcus.berkeley.edu/MANSWWW/MANSWWW.html
12. Freeman, W.J.: A pseudo-equilibrium thermodynamic model of information processing in nonlinear brain dynamics. Neural Netw. **21** (2008) 257–265 http://repositories.cdlib.org/postprints/2781
13. Barrie, J.M., Freeman, W.J., Lenhart, M.: Modulation by discriminative training of spatial patterns of gamma EEG amplitude and phase in neocortex of rabbits. J. Neurophysiol. **76** (1996) 520–539.
14. Freeman, W.J., Baird, B.: Relation of olfactory EEG to behavior: spatial analysis. Behav. Neurosci. **101** (1987) 393–408.
15. Freeman, W.J., O'Nuillain, S., Rodriguez, J.: Simulating cortical background activity at rest with filtered noise. J. Integr. Neurosci. **7**(3) (2008) 337–344.
16. Freeman, W.J.: Hilbert transform for brain waves. Scholarpedia **2**(1) (2007) 1338.
17. Percival, D.B., Walden, A.T.: Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. Cambridge: Cambridge University Press (1993).
18. Freeman, W.J., Erwin, H.: Freeman K-set. Scholarpedia **3**(2) (2008) 3238.
19. Freeman, W.J., Breakspear, M.: Scale-free neocortical dynamics. Scholarpedia **2**(2) (2007) 1357.
20. Freeman, W.J., Zhai, J.: Simulated power spectral density (PSD) of background electrocorticogram (ECoG). Cogn. Neurodyn. **3**(1) (2009) 97–103.
21. Miller, K.J., Sorenson, L.B., Ojemann, J.G., den Nijs, M.: Power-law scaling in the brain surface electric potential. PLOS Comp. Biol. **5**(12) (2009) e1000609 doi:10.1371/journal.pcbi.1000609.
22. Freeman, W.J., Rogers, L.J., Holmes, M.D., Silbergeld, D.L.: Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. J. Neurosci. Methods **95** (2000) 111–121.
23. http://sulcus.berkeley.edu/Null_Spikes_Movies
24. Freeman, W.J., Kozma, R.: Freeman's mass action. Scholarpedia **5**(1) (2010) 8040.
25. Rice, S.O.: Mathematical Analysis of Random Noise – and Appendixes – Technical Publications Monograph B-1589. New York, NY: Bell Telephone Labs Inc. (1950) Sect. 3.8., p. 90, equ. 3.8–15.
26. Haken, H.: Synergetic Computers & Cognition. Berlin: Springer (1991).
27. Pribram, K.: Holonomic brain theory. Scholarpedia **2**(5) (2007) 2735.
28. Ruiz, Y., Li, G., Freeman, W.F., Gonzalez, E.: Detecting stable phase structures on EEG signals to classify brain activity amplitude patterns. J. Zhejiang Univ. **10**(10) (2009) 1483–1491.
29. Vitiello, G.: My Double Unveiled. Amsterdam: John Benjamins (2001).
30. Freeman, W.J.: On the problem of anomalous dispersion in chaoto-chaotic phase transitions of neural masses, and its significance for the management of perceptual information in brains. In Haken, H., Stadler, M., eds.: Synergetics of Cognition, vol. 45. Berlin: Springer (1990) 126–143.
31. Izhikevich, E., Edelman, G.: Large-scale model of mammalian thalamocortical systems. PNAS **105**(9) (2008) 3593–3598.

Nonequilibrium Behavior in Neural Networks: Criticality and Optimal Performance

J.J. Torres, S. Johnson, J.F. Mejias, S. de Franciscis, and J. Marro

Abstract We present a general theory which allows one to study the effects on emergent, cooperative behavior of a complex interplay between different dynamic processes that occur in actual systems at the neuron, synapse and network levels. We consider synaptic changes at different time scales from less than the millisecond to the scale of learning, and the possibility of finding a fraction of silent neurons. For some limits of interest, the fixed-point solutions or *memories* then loose stability and the system shows enhancement of its response to changing external stimuli for particular network topologies and *dynamical* memories. We observe at the edge of chaos that the network activity becomes critical in the sense that the relevant quantities show non-trivial, power–law distributions. We also describe the effect of activity–dependent synaptic processes on the network storage capacity.

Keywords Unstable dynamics · Activity-dependent synaptic processes · Dynamical memories · Optimum network topology · Criticality

1 Introduction

Recent research in neuroscience including both in vivo and in vitro experiments have demonstrated that synapses are more than simple communication lines among neurons, and that many different dynamic processes taking place in the synapses can influence and even determine different type of information processing in the brain [1]. Some of these mechanisms can occur on different time scales. For instance, on a time scale longer than the second (say days or years), synapses can be modified due to *learning*. This has been widely studied within a general theory of learning in attractor neural networks [2]. In addition to this, it has been described that fast synaptic fluctuations coupled with other mechanisms during the transmission

J.J. Torres (✉)
Institute Carlos I for Theoretical and Computational Physics, University of Granada, E-18071 Granada, Spain
e-mail: jtorres@onsager.ugr.es

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_89,

of information seem to determine a large variety of computations in the brain [3, 4]. These fluctuations occur on very short (less than the millisecond) temporal scales, and they seem to have different causes. For instance, the stochasticity of the opening and closing of the neurotransmitter vesicles, variation in the postsynaptic response along the dendritic tree, which in turn has several sources (e.g., variations of the glutamate concentration in the synaptic cleft) and differences in the power released from different locations on the active zone of the synapses [5]. Together with these fast synaptic changes, it has been reported that the postsynaptic response is also affected by short-time activity-dependent mechanisms which can decrease or increase the amount of available neurotransmitter and, consequently, the postsynaptic response is either depressed or facilitated [1, 6]. This type of synaptic plasticity is believed to be fundamental for the development and adaptation of the nervous system, and to be at the base of higher brain functions such as learning and memor.

In this paper we present an attempt towards a theoretical framework to study systematically the influence of synaptic changes on the collective properties of a neural network, where the network topology itself is also a variable to be considered. In particular, of special interest is to understand how these synaptic mechanisms for different network topologies affect the fixed points of the neural activity and their stability, which concerns memory, recall processes and sensibility to external stimuli.

2 Model and Results

Let us consider N neurons – for simplicity assumed binary so that configurations are $\mathbf{S} \equiv \{s_i = \pm 1; i = 1, \ldots, N\}$ – connected by synapses of intensity $w_{ij} = \overline{w}_{ij} z_j \;\; \forall i,j$. Here, $\overline{w}_{ij} \equiv 1/N \sum_{\mu=1}^{M} \Xi_i^\mu \Xi_j^\mu$ are fixed and determined in a previous *learning* process in which the M patterns of neural activity $\Xi^\mu \equiv \{\Xi_i^\mu = \pm 1; i = 1, \ldots, N\}$, $\mu = 1 \ldots M$, are *stored*. $\overline{w}_{ij}$ represents the maximal averaged synaptic conductance between the presynaptic neuron j and the postsynaptic neuron i, while, $z_j \in \mathbb{R}$ is a stochastic variable that influences these conductances and accounts for other synaptic dynamics than those associated to learning. For fixed $\mathbf{W} \equiv \{\overline{w}_{ij}\}$, the network state $\mathbf{A} = (\mathbf{S}, \mathbf{Z} \equiv \{z_i\})$ follows the probabilistic dynamics

$$\frac{\partial P_t(\mathbf{A})}{\partial t} = \sum_{\mathbf{A}'} \left[P_t(\mathbf{A}') c(\mathbf{A}' \to \mathbf{A}) - P_t(\mathbf{A}) c(\mathbf{A} \to \mathbf{A}') \right], \tag{1}$$

where $c(\mathbf{A} \to \mathbf{A}') = p\, c^{\mathbf{Z}}(\mathbf{S} \to \mathbf{S}')\, \delta_{\mathbf{Z},\mathbf{Z}'} + (1-p)\, c^{\mathbf{S}}(\mathbf{Z} \to \mathbf{Z}')\, \delta_{\mathbf{S},\mathbf{S}'}$ [7]. This amounts to assume that neurons ($\mathbf{S}$) change stochastically in time competing with a noisy dynamics of synapses ($\mathbf{Z}$), the latter with an *a priory* relative weight of $(1-p)/p$ [8].

For $p = 1$, the model reduces to the Hopfield case, in which synapses are *quenched*, i.e., z_i is constant and independent of i, e.g., $z = 1$. This limit has been widely studied in the last decades [2]. More interesting is the case of $p \to 0$, which

describes fast synaptic fluctuations. In this limit, one can uncouple the stochastic dynamics for neurons ($\mathbf{S}$) and the synaptic noise ($\mathbf{Z}$) using standard techniques [8]. It follows that neurons evolve as in the presence of a steady distribution for the noise $\mathbf{Z}$: If we write $P_t(\mathbf{A}) = P_t(\mathbf{Z}|\mathbf{S})P_t(\mathbf{S})$, where $P_t(\mathbf{Z}|\mathbf{S})$ stands for the conditional probability of $\mathbf{Z}$ given $\mathbf{S}$, one obtains from (1), after rescaling time $tp \to t$ and summing over $\mathbf{Z}$, that

$$\frac{\partial P_t(\mathbf{S})}{\partial t} = \sum_{\mathbf{S}'} \left\{ P_t(\mathbf{S}')\bar{c}[\mathbf{S}' \to \mathbf{S}] - P_t(\mathbf{S})\bar{c}[\mathbf{S} \to \mathbf{S}'] \right\}. \tag{2}$$

Here, $\bar{c}[\mathbf{S} \to \mathbf{S}'] \equiv \sum_{\mathbf{Z}} P^{\rm st}(\mathbf{Z}|\mathbf{S})\, c^{\mathbf{Z}}[\mathbf{S} \to \mathbf{S}']$, and the stationary distribution for the noise is

$$P^{\rm st}(\mathbf{Z}|\mathbf{S}) = \frac{\sum_{\mathbf{Z}} c^{\mathbf{S}}[\mathbf{Z}' \to \mathbf{Z}]\, P^{\rm st}(\mathbf{Z}'|\mathbf{S})}{\sum_{\mathbf{Z}} c^{\mathbf{S}}[\mathbf{Z} \to \mathbf{Z}']}. \tag{3}$$

This expression involves an assumption on how synaptic noise depends on the overall neural activity. An interesting particular situation is to assume activity-dependent synaptic *noise* consistent with short-term synaptic depression and/or facilitation [6, 9]. That is, let us assume that $P^{\rm st}(\mathbf{Z}|\mathbf{S}) = \prod_j P(z_j|\mathbf{S})$ with

$$P(z_j|\mathbf{S}) = \zeta\,(\mathbf{m})\;\delta(z_j - \Phi) + [1 - \zeta\,(\mathbf{m})]\;\delta(z_j - 1). \tag{4}$$

Here, $\mathbf{m} = \mathbf{m}(\mathbf{S}) \equiv \left(m^1(\mathbf{S}), \ldots, m^M(\mathbf{S})\right)$ is the M-dimensional overlap vector, $m^\mu = N^{-1} \sum_i \Xi_i s_i$, and $\zeta\,(\mathbf{m})$ stands for a function of $\mathbf{m}$ to be determined. With this choice, the average over the distribution (4) of the noise variable is $\overline{z_j} \equiv \int z_j P(z_j|\mathbf{S}) dz_j = 1 - (1 - \Phi)\zeta\,(\mathbf{m})$ and the variance is $\sigma_z^2 = (1 - \Phi)^2 \zeta\,(\mathbf{m})$ $[1 - \zeta\,(\mathbf{m})]$. Note that these two quantities depend on time for $\Phi \neq 1$ through the overlap vector $\mathbf{m}$, which is a measure of the activity of the network. Moreover, the depression/facilitation effect in (4), namely $z_j = \Phi > 0$ ($\Phi \neq 1$), depends through the probability $\zeta\,(\mathbf{m})$ on the overlap vector, which is related to the net current arriving to postsynaptic neurons. Consequently, the non–local choice (4) introduces non–trivial correlations between synaptic noise and neural activity. One has a depressing (facilitating) effect for $\Phi < (>)1$, and the trivial case $\Phi = 1$ corresponds to the Hopfield model with quenched synapses. It is remarkable that, although the fast noise dynamics occurs at a very small time scale, the depressing or facilitating mechanism occurs on the time scale of the neural activity –via the coupling with the overlap vector through the function $\zeta\,(\mathbf{m})$.

The general model described by Eqs. (1–4) can be easily generalized to other cases of interest such as the possibility of having silent nodes in the network (every time the activity is updated). For instance, one may assume that the transition probabilities have the form

$$c^{\mathbf{Z}}[\mathbf{S} \to \mathbf{S}'] = \sum_{\mathbf{x}} p_n(\mathbf{x}) \prod_{i|x_i=1} \tau_n(s_i \to s_i'; \mathbf{Z}) \prod_{i|x_i=0} \delta_{s_i, s_i'}. \tag{5}$$

Here, $\mathbf{x}$ is an operational set of binary indexes fixed to 1 at n sites chosen at each time according to distribution $p_n(\mathbf{x})$, and fixed to zero at the other $N - n$ sites. The choice (5) simply states that one (only) updates simultaneously the selected n nodes. We also assume that the elementary rate $\tau_n(s_i \to s_i'; \mathbf{Z})$ depends on the factor $\beta s_i h_i$ where $\beta = T^{-1}$ is the inverse of the temperature and $h_i(\mathbf{S}, \mathbf{Z}) \equiv \sum_j \epsilon_{ij} \overline{w}_{ij} z_j s_j$ is the local field or synaptic current a particular neuron is receiving from its neighbors. Here $\epsilon_{ij} = \{1, 0\}$ is the adjacency or connectivity matrix, which describes the existence or not of interaction between i and j. This allows for the consideration of different network topologies. For any non-trivial topology, is convenient to consider the *local* overlap $m_j^\mu = \langle k \rangle^{-1} \sum_i \Xi_i s_i$, where $\langle k \rangle$ is the mean neuron degree or number of neighbors, which equals the *global* one only when all neurons are connected to each other.

A main result is that the system shows quite qualitatively different behavior depending on the value of T and $\rho \equiv n/N$. In particular, one easily observes the familiar Hopfield phases of complete disorder and one in which the system may recover one of the stored patterns. More intriguing are cases such as those in Fig. 1, namely, dynamic phases in which the network activity chaotically switches among the stored patterns and antipatterns (i.e., negatives of the stored patterns).

Concerning topology, a given neuron i in biological networks is seldom or never connected to the rest, but rather only to a relatively small subset of neighbors – of cardinal $k_i = \sum_j \epsilon_{ij}$, which is i's *degree*. Even in the case of the worm *C. elegans*, with only about 300 neurons, it turns out that there is great disparity in the values of k, ranging from just one or two to a large portion of the network. Although little is yet known about the precise architecture of animal brains, a first approximation we can consider for neural systems – as is often done in other networks – is to assume that the synapses are placed randomly between the neurons, and focus only on the degree distribution, $p(k)$. In this approximation, known as the *configurations ensemble*, the expected value of ϵ_{ij}, given k_i and k_j, is $\overline{\epsilon_{ij}} = k_i k_j / (\langle k \rangle N)$. We will be particularly interested in *scale-free* distributions, $p(k) \sim k^{-\gamma}$, which appear to be quite ubiquitous in nature.

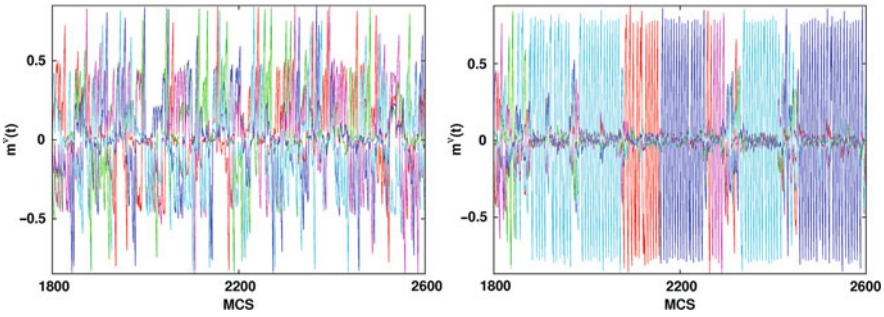


Fig. 1 Two of the different types of dynamic behaviour exhibited by the nonequilibrium neural network showing chaotic switching from one attractor to the other. Parameters are $N = 1,600$, $M = 5$, $T = 0.01$, $\Phi = -0.8$ and $\rho = 0.45(0.5)$ for the *left* (*right*) simulation

Standard mean-field analysis for $M = 1$ (a single stored pattern and its antipattern) [10] shows that there is a second order phase transition, from the memory phase to the chaotic one, with increasing temperature. The critical value for $\rho = 1$ is $T_c = \langle k^2 \rangle / (\langle k \rangle N)$, where the averages $\langle \cdot \rangle$ are over $p(k)$. Setting $T = 0$, we find there is also a transition, with decreasing Φ, from the memory phase in which either of the fixed points $m = \pm 1$ is stable, to one in which the system jumps chaotically between the two. The critical value of Φ at which this occurs is $\Phi_0 = 1 - \langle k \rangle^3 \langle k^3 \rangle^{-1}$. MC simulations show that, for non-zero T, chaotic windows open for $\Phi \lesssim \Phi_0$. They also tell us that these results are robust for larger values of M and qualitatively similar for $\rho < 1$.

This *edge of chaos* is particularly interesting, since it has been shown that this kind of transition can be optimal for certain magnitudes such as computational capacity [11] and dynamic range of sensitivity to stimuli [12]. To illustrate how this is also the case here, we store a set of M patterns and then *show* the system a randomly chosen one every certain number of time steps. This is done by changing the field at each node for just one MCS: $h_i \to h_i + \delta \xi_i^{\nu}$, as if a signal of intensity δ had been received. In general, when in the ordered phase, a high δ is required to destabilize the current pattern. However, close to the transition to chaos, the system becomes more sensitive and reacts appropriately. Within the chaotic regime, however, it does not then retain the pattern shown. We quantify performance as the temporal average $\langle m^{\nu} \rangle_{time}$, where ν is the pattern last shown. This is displayed in Fig. 2a. The figure shows that the performance exhibits a peak corresponding to the optimal scale-free topology, that for the case of low depression ($\Phi \to 1$) occurs around $p(k) \sim k^{-2}$. This is a consequence of the particular dependence of the critical parameters T_c and Φ_c with the moments of the degree distribution: the more heterogeneous the network, the more robust the system will be to thermal fluctuations, and yet the less synaptic depression will be required to place the system at the edge of chaos. This optimization effect also is reflected by the emergence of power-law distributions of permanence times around the attractors, when the system in near the edge of chaos, as is depicted in Fig. 2b. Here, $P(\Delta\tau)$ is the probability that, throughout the time interval $(t', t' + \Delta\tau)$, a local field $h_i(t)$ is fluctuating above a baseline $h_0 = 0.1$ or below $h_0 = -0.1$, and averaged over all sites in the network. Optimal scale-free network topologies with exponent -2 could be related to the existence of a functional scale-free topology of the human brain during cognitive tasks with the same exponent [13]. It seems plausible, therefore, that the brain may acquire this kind of distribution – either structurally or functionally – in order to maximize its performance.

Finally we also report on the effect of synaptic processes coupled with network activity on maximum storage capacity of the network [14] via a phenomenological model of activity-dependent synapses (see [15] for details) which involves a competition between facilitating and depressing synaptic mechanisms. This model can be studied using our general theoretical framework assuming $P(z_j/\mathbf{S}) = \delta(z_j - \Phi_j(\mathbf{S}))$ with $\Phi_j(\mathbf{S})) = D_j(t, s_j) F_j(t, s_j)$ and $D_j(t, s_j)$, $F_j(t, s_j)$ representing dynamic variables for the depression and facilitation mechanisms varying in time-scales τ_{rec} and τ_{fac}, respectively. Maximum storage capacity is computed at $T = 0$ and in the

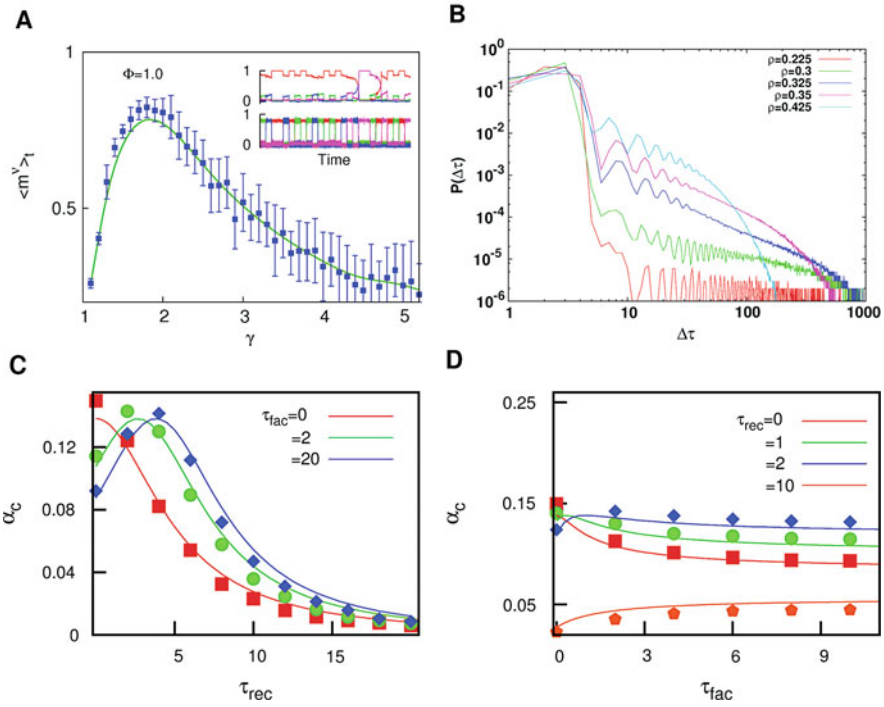


Fig. 2 (**a**) Network performance against γ for scale-free random topologies, with $\Phi = 1.0$ (static-synapses limit). Averages over 20 realisations, with stimulation every 50 MCS for 2,000 MCS. Other parameters are $\delta = 5$, $M = 4$, $T = 2/N$, $\langle k \rangle = 20$, $N = 1,600$, $\rho = 1$. Inset: sections of typical time series of m^{ν}, with $\nu = 1, ..., 4$ represented as different colours, for $\gamma = 4$ and Φ (*top*) and $\Phi = 0.6$ (*bottom*). (**b**) Emergence of criticality in the model near the edge of chaos. Other parameter were $N = 800$, $M = 10$, $T = 0$, $\Phi = -0.75$ and $\Delta = 0.1$. (**c**) Critical storage capacity α_c as a function of τ_{rec}, for different values of τ_{fac}. (**d**) Critical storage capacity α_c as a function of τ_{fac}, for different values of τ_{rec}. In both panels, the results from numerical simulations (symbols) with ($N = 3,000$) are supported by mean field predictions (*lines*)

memory phase. In practice, this implies considering only the steady state of D_j and F_j as a function of s_j and including them in the local field h_i. When the number of stored patterns is increased, the interference between patterns makes the attractors associated with these patterns loose their stability. Then, the maximum number of patterns, relative to the network size, namely $\alpha_c = M_c/N$, that the network can store and retrieve without interference is the maximum storage capacity. The behaviour of α_c as a function of the relevant synaptic parameters is depicted in Fig. 2c, d. In general, one has a non-monotonic dependence of storage capacity as a consequence of the competition of different synaptic processes with, a priori, opposite effects on the stability of attractors (e. g., depression tries to destabilize the memories whereas facilitation tries to keep the activity of the network in one particular attractor). This fact implies that activity-dependent synaptic changes are not only convenient for dynamic processing of information in actual neurons, as we have seen above, but

an optimal balance between depression and facilitation effects is also necessary to have neural networks with good retrieval properties.

Acknowledgements Financed by projects JA FQM–01505 and MEC–FEDER FIS2009–08451.

References

1. Abbott, L.F., Regehr, W.G.: Synaptic computation. Nature. **431**, (2004) 796–803.
2. Peretto, P.: An Introduction to the Modeling of Neural Networks. Cambridge: Cambridge University Press (1992).
3. Allen, C. Stevens, C.F.: An evaluation of causes for unreliability of synaptic transmission. Proc. Natl. Acad. Sci. USA **91** (1994) 10380–10383.
4. Zador, A.: Impact of synaptic unreliability on the information transmitted by spiking neurons. J. Neurophysiol. **79**, (1998) 1219–1229.
5. Franks, K.M., Stevens C.F., Sejnowski, T.J.: Independent sources of quantal variability at single glutamatergic synapses. J. Neurosci. **23** (2003) 3186–3195.
6. Tsodyks, M.V., Pawelzik, K., Markram, H.: Neural networks with dynamic synapses. Neural Comp. **10** (1998) 821–835.
7. Torres, J.J., Garrido P.L., Marro, J.: Neural networks with fast time-variation of synapses. J. Phys. A: Math. Gen. **30** (1997) 7801–7816.
8. Marro, J., Dickman, R.: Nonequilibrium Phase Transitions in Lattice Models. Cambridge: Cambridge University Press (1999).
9. Pantic, L., Torres, J.J., Kappen, H.J., Gielen S.C.A.M.: Associative memory with dynamic synapses. Neural Comp. **14** (2002) 2903–2923.
10. Johnson, S., Marro, J., Torres, J.J.: Functional optimization in complex excitable networks. EPL **83** (2008) 46006.
11. Bertschinger, N., Natschläger, T.: Real-time computation at the edge of chaos in recurrent neural networks. Neural Comp. **16** (2004) 1413–1436.
12. Assis, V.R.V., Copelli, M.: Dynamic range of hypercubic stochastic excitable media Phys. Rev. E **77** (2008) 011923.
13. Eguíluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., Apkarian, A.V.: Scale-free brain functional networks. Phys. Rev. Lett. **95** (2005) 018102.
14. Torres, J.J., Pantic, L., Kappen, H.J.: Storage capacity of attractor neural networks with depressing synapses. Phys. Rev. E. **66** (2002) 061910.
15. Mejias, J.F., Torres, J.J.: Maximum memory capacity on neural networks with short-term synaptic depression and facilitation. Neural Comput. **21** (2009) 851–871.

Part X
Neural Computing

A Hierarchial Model for Visual Perception

Bolei Zhou and Liqing Zhang

Abstract This paper proposes a computational model for visual perception: the visual pathway is considered as a functional process of dimensionality reduction for input data, that is, to search for the *intrinsic* dimensionality of natural scene images. This model is hierarchically constructed, and final leads to the formation of a low-dimensional space called *perceptual manifold*. Further analysis of the perceptual manifold reveals that scene images which share similar perceptual similarities stay nearby in the manifold space, and the dimensions of the space could describe the spatial layout of scenes, which are like the degree of naturalness, openness supervised trained in (Oliva, A., and Torralba, A. Int. J. Comput. Vis. 42 (2001) 145–175). Moreover, the implementation of scene retrieval task validates the topographic property of the perceptual manifold space.

Keywords Sparse coding · Hierarchical model · Computational vision

1 Introduction

From photoreceptor of retina to the unified scene perception in higher level cortex, the transformations of input signal in visual cortex are very complicated. The neurophysiological studies [1] indicate that through the hierarchical processing of information flow in visual cortex, the extremely high-dimensional input signal is gradually represented by fewer active neurons, which is believed to achieve a sparse coding [2] or efficient coding [3].

Meanwhile, the efficient coding theory proposed by Horace Barlow in 1961 gradually becomes the theoretical principles for nervous system and explain away much neuronal behaviors. One of the key predictions of efficient coding theory is that the visual cortex relies on the environmental statistics to encode and represent the signal [4], that is to say, the neuronal representation is closely related to the intrinsic

B. Zhou (✉)
MOE-Microsoft Laboratory for Intelligent Computing and Intelligent Systems, Department of Biomedical Engineering Shanghai Jiao Tong University, Shanghai, China
e-mail: zhoubolei@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_90,

properties of the input signal. Moreover, studies on the natural image statistics show that the natural images are usually embedded in a relatively low dimensional manifold of pixel space [5], and there is a large amount of information redundancy within the natural images. According to the efficient coding theory, it is assumed that the structure of neural computations should fully adapt to extract the intrinsic low dimensionality of natural image to form the unified scene perception, in spite of the extremely high-dimensional raw sensory input from the retina [1].

Under the efficient coding theory and natural image statistics, we propose a computational model for visual perception: the visual pathway is considered as a functional process of dimensionality reduction for input signal, that is, to remove the information redundancy and to search for the *intrinsic* dimensionality of natural scene images. By pooling together the activity of local low-level feature detectors across large regions of the visual field, we build the population feature representation which is the statistical summary of the input scene image. Then, thousands of population feature representations of scene images are extracted, and to be mapped *unsupervised* to a low-dimensional space called perceptual manifold. Further analysis of this manifold reveal that scene images which share similar perceptual similarity stay nearby in the manifold space, and the dimensions of the manifold could describe continuous changes within the spatial layout of scenes, which are similar to the degree of naturalness and openness supervised trained in [6]. In addition, the implementation of scene retrieval task validates the topographic property of the perceptual manifold space. In the following section, the Perceptual Manifold model is extended in detail.

2 The Hierarchical Model

One of the fundamental properties of visual cortex concerns the ability to integrate the local components of a visual image into a unified perception [1]. Based on the neurobiological mechanism of hierarchical signal processing in visual system, we build a computational architecture called Perceptual Manifold. The architecture includes three hierarchical layers: (1) local feature encoding, (2) population feature encoding and (3) perceptual manifold embedding (refer to Fig. 1a).

2.1 Local Sparse Feature Encoding

Experimental studies have shown that the receptive fields of simple cells in the primary visual cortex produce a sparse representation of input signal [4]. The standard efficient coding method [7] assumes that the image patch is transformed by a set of linear filters w_i to output response s_i. In matrix form,

$$\mathbf{s} = \mathbf{W}\mathbf{x} \tag{1}$$

Or equivalently in terms of a basis set, $\mathbf{x}=\mathbf{A}\mathbf{s}=\mathbf{W}^{-1}\mathbf{s}$. Then, the filter response s_i are assumed to be statistically independent,

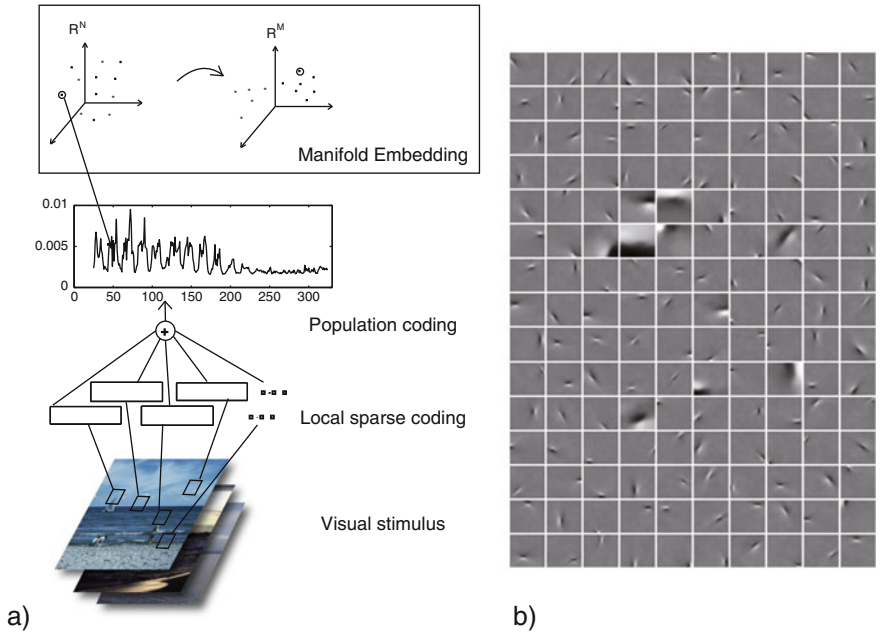


Fig. 1 **(a)** Schematic diagram of the Perceptual Manifold model. There are three hierarchical layers of sensory computation before the final formation of the perceptual space. **(b)** A subset of filter basis **W** trained from natural image patches, they resemble receptive fields of the simple cells in V1

$$p(\mathbf{s}) = \prod_i p(s_i) \propto exp(-\sum_i |s_i|) \tag{2}$$

Specifically, 150,000 20 × 20 gray image patches from natural scenes are used in training. A set of 384 20×20 basis function is obtained. These basis functions resemble the receptive field properties of simple cells, i.e., they are spatially localized, oriented, and band-pass in different spatial frequency bands (Fig. 1b). Let $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{384}]$ be the filter function. A vectorized image patch $\mathbf{x}$ can be decomposed into those statistically independent bases, in which only a small portion of bases are activated at one time. They are used as the first layer of local feature extraction in the architecture.

2.2 Population Feature Encoding

Higher processing stages are not influenced by single V1 neuron alone but by the activity of a larger population [8]. The neurophysiological study [1] implies that on the population level, a linear pooling mechanism might be used by the visual cortex to extract the global response of the stimulus. In view of this evidence, our computational model includes a second layer of population encoding to sum up feature's response over the local visual fields.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, ...]$ denote the sample matrix, where $\mathbf{x}_n$ is the vectorized image patch sampled from scene image. The population feature component for the i^{th} feature is:

$$p_i = \frac{\sum_n |\mathbf{w}_i^\top \mathbf{x}_n|}{\sum_i \sum_n |\mathbf{w}_i^\top \mathbf{x}_n|} \tag{3}$$

Thus, $\mathbf{p} = [p_1, p_2, ..., p_{384}]^\top$ is the population response of scene image.

2.3 Perceptual Space Embedding

Furthermore, neurophysiological studies have often found that the firing rate of each neuron in a population can be written as a smooth function of a small number of variables [8], which supports the idea that the population activity might be constrained to lie on a low-dimensional manifold. To search for the underlying meaningful dimensionality of percept, the Local Linear Embedding [9] is applied as the nonlinear dimensionality reduction method for thousands of the population feature responses of scene images:

First step: compute the weight W_{ij} that best linearly reconstructs $\mathbf{p}_i$ from its neighbor $\mathbf{p}_j$, minimizing:

$$\varepsilon(W) = \sum_i |\mathbf{p}_i - \sum_j W_{ij}\mathbf{p}_j|^2 \tag{4}$$

Second step: compute the low-dimensional embedding vectors $\mathbf{y}_i$ best reconstructed by W_{ij}, minimizing:

$$\phi(\mathbf{y}) = \sum_i |\mathbf{y}_i - \sum_j W_{ij}\mathbf{y}_j|^2 \tag{5}$$

The resulting embedding space $\mathbb{R}^M$ is called perceptual space, where vector $\mathbf{y} = [y_1, y_2, ..., y_M]^\top \in \mathbb{R}^M$. The topographic properties of this space would be analyzed in the following experiment section.

3 The Experiment

The dataset of natural images implemented here is from [10], which contains 3,890 images from 13 semantic categories of natural scenes, like coast, forest, etc. And image size is normalized as 128×128 pixels.

All 3,890 images are used in the embedding process. The dimensionality of manifold space M is tuned as 20, so that the space embedding is $\mathbb{R}^{384} \rightarrow \mathbb{R}^{20}$. Figure 2 shows the embedded sample points described by first three coordinates of the perceptual space, in which points of images from four categories are visualized, and

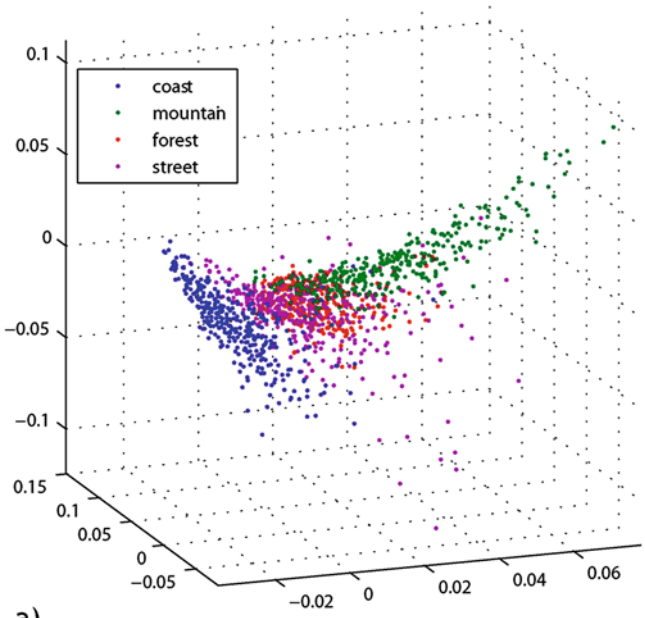


Fig. 2 **(a)** *Colored points* of scene images from four categories are visualized by the first three coordinates of the perceptual space, the geometric structure among those points of scene images is clearly nonlinear. **(b)** Some examples of target images with the sets of retrieved images. Despite the simplicity of the similarity metric used in the experiment, these pairs share close perceptual similarity (Colors will be viewed in online version)

are colored according to the images' scene category. We can clearly see the clustering and nonlinear geometric property of the data points in the space. Moreover, the image retrieval task as follows validates its property.

3.1 Image Retrieval

Image retrieval task is to retrieve the perceptually similar sets of images when given a target image. This task is implemented in our embedded perceptual space and achieves impressive results.

To show the topographic property of the perceptual space, image similarity metric is approximated simply as the Euclidean distance between the target i and retrieved ones j in the perceptual space, that is,

$$D^2(i,j) = \sum_{h=1}^{M} (y_{ih} - y_{jh})^2 \tag{6}$$

Fig. 2 shows examples of target images and retrieved sets of K least metric images, here K is 6. Despite the simplicity of the similarity metric used in the experiment, the set of retrieved images are very similar to the target image.

4 Discussion and Conclusion

The efficiency of the neural code depends both on the transformation that maps the input to the neural responses and on the statistics of the input signal [4].

In this paper, under the theory of efficient coding, a hierarchical framework is developed to explore the intrinsic dimensions of visual perception. By pooling together the activity of local low-level feature detectors across large region of the visual field, we build the population feature representation which is the statistical summary of the input scene image. Then, thousands of population feature representations of scene images are extracted, and mapped *unsupervised* to the low dimensional perceptual space. Analysis of the perceptual space reveals the topographic property that scene images with similar perceptual similarity stay nearby in the embedded space. Moreover, the implementation of scene retrieval task validates the topographic property of the perceptual space.

In addition, it is noteworthy that in the image retrieval task, most of the retrieved images belong to the same scene category of the target image hand-labored in [10]. This would imply the topographic property of the embedded perceptual space corresponds to the semantic concept of scene image, and this interesting property would be discussed in our further work.

Acknowledgements The work was supported by the National Basic Research Program of China (Grant No. 2005CB724301) and the Science and Technology Commission of Shanghai Municipality (Grant No. 08511501701).

References

1. Weliky, M., Fiser, J., Hunt, R.H., Wagner, D.N.: Coding of natural scenes in primary visual cortex. Neuron **37**(4) (February 2003) 703–718.
2. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**(6583) (June 1996) 607–609.
3. Olshausen, B.A., Field, D.J.: Sparse coding of sensory inputs. Curr. Opin. Neurobiol. **14**(4) (August 2004) 481–487.
4. Simoncelli, E.P., Olshausen, B.: Natural image statistics and neural representation. Ann. Rev. Neurosci. **24** (May 2001) 1193–1216.
5. Srivastava, A., Lee, A.B., Simoncelli, E.P., c. Zhu, S.: On advances in statistical modeling of natural images. J. Math. Imaging Vis. **18** (2003) 17–33.
6. Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. Int. J. Comput. Vis. **42** (2001) 145–175.
7. Bell, A.J., Sejnowski, T.J.: The "independent components" of natural scenes are edge filters. Vision Res. **37**(23) (December 1997) 3327–3338.
8. Averbeck, B.B., Latham, P.E., Pouget, A.: Neural correlations, population coding and computation. Nat. Rev. Neurosci. **7**(5) 358–366.
9. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**(5500) (December 2000) 2323–2326.
10. Li, F.F., Perona, P.: A bayesian hierarchical model for learning natural scene categories. In CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2. Washington, DC: IEEE Computer Society (2005) 524–531.

An Orientational Sensitive Vision Model Based on Biological Retina

Hui Wei and Yuxiang Jiang

Abstract This paper described a novel approach to construct a space-variant cell distribution according to biological retina. The model designed in this paper samples pictures with pseudo-random "photoreceptor cells" rather than with pixels. And such a sub-pixel sampling method radically differs from uniform pixel array version used in conventional computational vision. Moreover, a simple type of ganglion cells was also modeled to represent orientational stimulus in the visual field.

Keywords Space-variant · Sub-pixel sampling · Orientation sensitive

1 Introduction

As we known, a bunch of preeminent advantages of retina have been evolved for human to survive. Retina, as an outstretched part of central neural system, contributes a lot to image processing, especially to integration and segmentation [1–3]. On one hand, The photoreceptor cells highly packed in central foveal region provide a fine resolution of the object we focused. On the other hand, visual accuracy decreases on the large periphery, which probably results in reducing the energy consumption. Despite the ganglion cells enlarge their receptive fields along with the eccentricity, vision fields do not seem to be blurred. In addition, it even ensures the capacity of keeping alert the danger in the surroundings.

Pictures used in conventional digital image processing are usually sampled by a matrix-like pixel array, which is uniformly distributed. The output produced by these methods (involving Fourier Transformation and filter operations in most cases) would always be represented on the same scale as the original picture, which does not conduce to understanding the contents of pictures.

H. Wei (✉)
Cognitive Algorithm Model Laboratory, School of Computer Science, Fudan University, 200433, Shanghai P.R. China
e-mail: weihui@fudan.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_91,

Some attempts have been made to transform from the orthogonal coordinates to the log-polar [4, 5]. And the foveal region has to be considered separately since the serious problem of singularity at center. Another alternative is to divide the entire retina to several rings around center with different cell densities [6]. On the other hand, [7] presented a self-organized version of retina to calculate the position of each cell, which simulated the gradually change of density successfully. But it cannot be so adapted for various kinds of distributions (e.g. rods density distribution) easily, due to such restricts as complicated parameter adjustment and numerous iterations.

2 Retina Cells

Although a digital image consists of pixels, the model in this paper doesn't regard these pixels as the minimum sample units. Each pixel with [R, G, B] values is considered as a continuous color block. The whole picture, thus, should be interpreted as an assemblage of many rectangle color blocks. Each photoreceptor cell has a coordinate for its position on the retina. This enables the photoreceptor cells to sample an input picture in definite position, while the picture is being mapped onto retina model. Therefore, not all the pixels – color blocks – around the periphery are necessarily sampled by photoreceptor cells. However, the pixels inside the foveal region may be over sampled by more than one photoreceptor cell.

In order to reduce the internal representation of the cells' position, this paper restricts these coordinates to locate on a large *grid*, see Fig. 1. In addition, small grids (8×8) – called *patterns* – are defined as the minimum units making up the entire grid. Each one of the 64 positions on the pattern either has or has not an

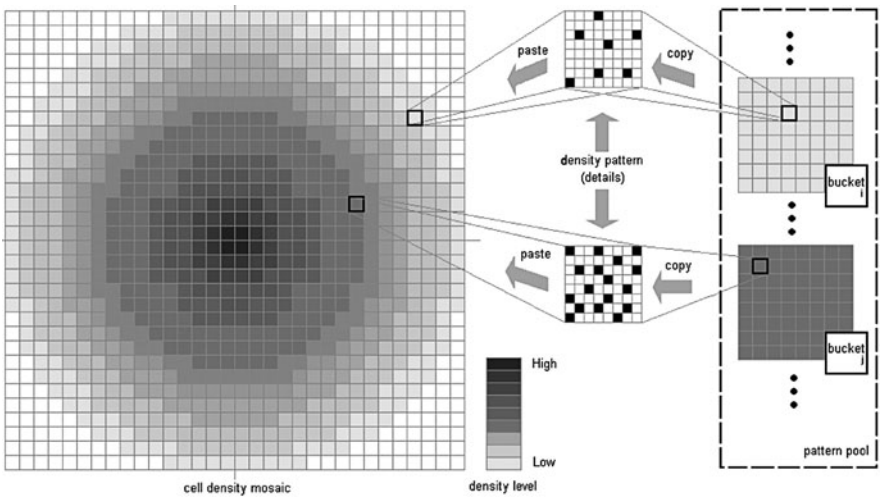


Fig. 1 Generate the cells using pattern pool. Each *black spot* on the 8×8 pattern represents a single cell. The model just generates multitudes of indices, each of which consists of a required *bucket-index* and a random *inner-bucket-index*, according to the density mosaic

existing cell, and the more cells on a pattern, the higher density the relative region has. Since the entire grid looks like a collage consisting of a lot of pattern "pieces", it is efficient to obtain a grid after adequate patterns being predefined – Patterns with the same number of existing cells are collected together into a *bucket*.

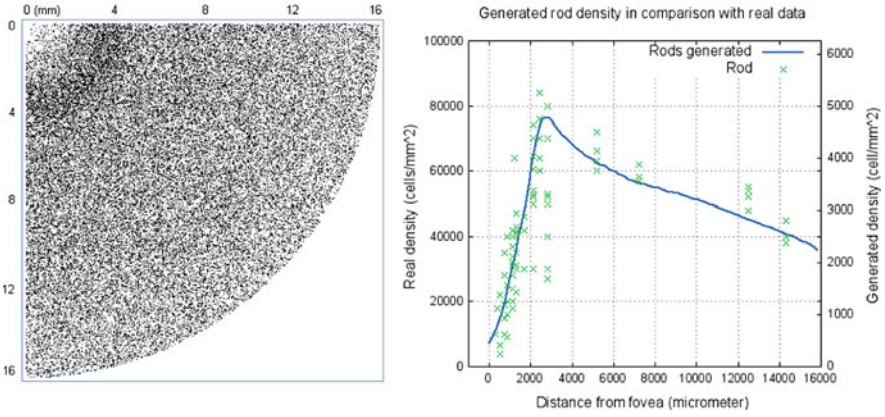
3 Receptive Fields

The ganglions (output layer) collect the responses of photoreceptors, and each cell holds a rectangular receptive field with the size depending on its eccentricity from the fovea (Fig. 3(a)). Ganglions receive inputs within this rectangular field and summate them with the *weights*, where positive values for excitations and negative ones for inhibitions. Eight types of receptive fields are used to form 8 orientation sensitive types of ganglions. And these kinds of ganglions contribute to the Magnocellular pathway and only receive rods' inputs in this simple model. The response of a ganglion is given by

$$Resp(n) = S(n) \sum_i rod_i W_i - E(n-1) \tag{1}$$

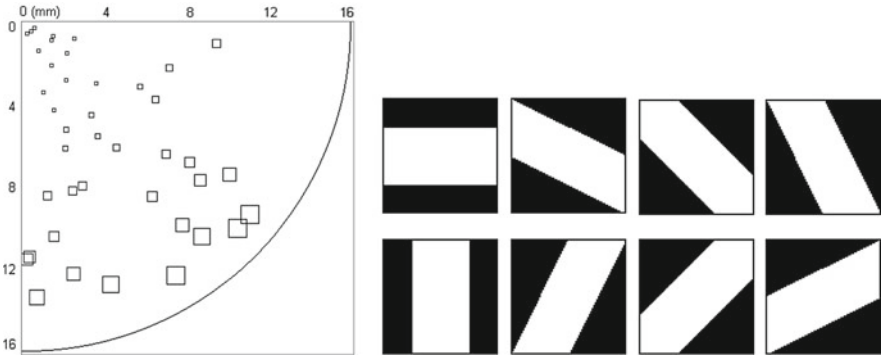
$$E(n) = \frac{(n-1)E(n-1) + Resp(n)}{n} \tag{2}$$

Here, W_i is the weight of each rod input, either excitation or inhibition, and $E(n-1)$ is the expectation of the response of $n-1$ pictures, which could be produced after dealing with $n-1$ training pictures by recursion, while $E(0) = 0$. $S(n)$ is the standardize scaler used to adjust the raw response.



(a) (b)

Fig. 2 *Right-down* quadrant of the rods generated, and in comparison with the real data [8]



(a) various RF size (b) 8 orientations

Fig. 3 **(a)** The receptive field size of ganglions' varies from fovea to periphery, and *bright regions* in **(b)** represent the excitations while *dark regions* for inhibitions

4 Results

Here is an example picture processed using the orientation sensitive model. The positive response of each ganglion has been rescaled to [0, 100], therefore, the response gives 100 only if all the positive weight inputs reach their maximum value and the negative region inside the receptive field receive no input.

For a nature picture, most ganglions are not "fired" sice the low response (Table 1). In other words, the artificial retina doesn't need to represent most of the ganglions, which effectively reduce the complexity and work load of the system.

Table 1 Statistics of the ganglion responses of the example above

Response intensity	1–2	2–3	3–4	4–5	5–6	>6
Number of ganglions	1189	523	264	146	112	473
Percentage (%)	4.47	2.00	0.99	0.62	0.42	1.78

Fig. 4 **(b)** is the response image of rods layer after sample the source picture. Each small *red bar* in **(c)** shows a single "fired" ganglion whose response exceeds the threshold, which was set as 5 in the case, and the longer a bar is displayed, the larger size its receptive field has (Colors will be viewed in online version)

The grass region in the picture can only trigger a few ganglions (Fig. 4(c)) to fire with a threshold fixed on 5.

5 Discussions

This paper shows an implementation of the orientation sensitive vision pathway based on space-variant foveated retina. As the main idea of this model, nonuniform sampling aims to reduce the fire number of ganglions, therefore, reduce the energy consumption. Yet, other cells, like cones, also contribute in the orientation sensitive pathway, and we have to take them into account in our further research works. It is important to notice that the imprecisely sampled information on periphery helps, indeed, a lot to understand the picture on a global scale. However, the Parvocellular and Magnocellular pathway of human vision, which account for over 80% of the ganglion outputs [9], function efficiently in classifying information according to colors, shapes, orientations and motions etc. [10]. The author believe that the study on such parallel pathway provide us a novel way to look into how does the retina represent the light information and why does this representation work so well.

Acknowledgements This work was supported by China-UK Science Networks Program 2008, co- founded by Chinese Ministry of Education and UK Office of Science and Innova- tion. In addition, it was supported by 973 Program (Project No. 2010CB327900) and by the key Sci-Tech program (Project No. 0851151703) of Shanghai Science and Technology Commission. We also appreciated the guidance given by Prof. Shou, Tiande from Department of Biophysics and Research Center for Brain Science, Fudan University.

References

1. Kolb, H.: How the Retina works. Am. Sci. 91 (2003) 28–35.
2. Masland, R.H.: The fundamental plan of the retina. Nature 4 (2001) 877–886.
3. Bloomfield, S.A., Dacheux, R.F.: Rod vision: pathways and processing in the Mammalian Retina. Prog. Retin. Eye Res. 20 (2001) 351–384.
4. Shah, S., Levine, M.D.: Visual information processing in Primate Cone Pathway—Part I: A Model. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 26 (1996a) 259–274.
5. Shah, S., Levine, M.D.: Visual Information processing in Primate Cone Pathway—Part II: Experiments. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 26 (1996b) 275–289.
6. Wei, H., Guan, X.: A computational vision model based on the physiological retina and its multi-task balance property. International conference on Image Processing, Computer Vision and Pattern Recognition. (2008) 446–452.
7. Balasuriya, S., Siebert, P.: A biologically inspired computational vision front-end based on a self-organised pseudo-randomly tessellated artificial retina. Proceeding of International Joint Conference on Neural Networks 3 (2005) 3069–3074.
8. Jonas, J.B., Schneider, U. et al.: Count and density of human retinal photoreceptors. Graefe's Arch. Clin. Exp. Ophthalmol. 230 (1992) 505–510.
9. Lee, B.B.: Receptive field structure in the Primate Retina. Vision Res. 36 (1996) 634–644.
10. Croner, L.J., Kaplan, E.: Receptive fields of P and M Ganglion cells across the Primate Retina. Vision Res. 35 (1995) 7–24.

Recurrent Neural Networks for Local Model Prediction

Aymen Cherif and Romuald Boné

Abstract "Local models" (Walter, J., et al. International Joint Conference on Neural Networks, vol. 1. (1990) 589–594), consists on dividing the data into homogeneous clusters by Vector Quantization (VQ (Gray, R. M., and Neuhoff, D.L. IEEE Trans. Inf. Theory 44(6) (1998) 2325–2383)) to simplify the prediction task on each cluster and mostly inspired from the Self-Organizing Maps algorithm (SOM (Kohonen, T. Self-Organization and associative memory, 3rd edn. (1989))). Since recurrent neural networks have demonstrated in many times a better results and specially for chaotic time series (Boné, R. Recurrent Neural Networks for Time Series Forecasting. (2000)), we propose in this paper a method to use the Recurrent Neural Networks in the local approach.

Keywords Time series prediction · Recurrent neural network · Self organizing map · Back propagation through time · Local approaches

1 Introduction

Time series forecasting is a widely treated problem which can be found in many disciplines. Researchers coming from various fields have addressed this question, for example in finance, mathematics, physics or machine learning. With this emphasis, many models and algorithms have been proposed. Starting by statistical methods [1], Multi-Layer Perceptron (MLP [2]) and Recurrent Neural Networks (RNN [2]). In addition to these methods, many researchers have proposed methods and techniques to push the limits of prediction error. A used technique in this last decade is the "local models" [3]. It consists on dividing the data into homogeneous clusters by Vector Quantization (VQ [4]) to simplify the prediction task on each cluster and mostly inspired from the Self-Organizing Maps algorithm (SOM [5]).

Since recurrent neural networks have demonstrated in many times a better results and specially for chaotic time series [6], we propose in this paper a method to use

A. Cherif (✉)
e-mail: cherif.aymen@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_92,

the Recurrent Neural Networks in the local approach. In the first section, a brief description of the previous work on the local approach is presented. In the second section, we will describe the used method in this paper. Then some tests will be presented and discussed. At last we will conclude with perspectives.

2 Vector Quantization and Forecasting

The Vector Quantization is the first step used in the local approach in order to divide data into clusters. The most popular and used algorithm is the SOM [5]. When used in local approach, the output layer can be changed from a 2D Layer to a linear Layer or even a circular one [7]. However, the original SOM algorithm does not take into account the temporal propriety (sequentiality) of the time series. A first variant is the Temporal Kohonen Map (TKM [8]), which performs a leaky integration in each output neuron. with these "leaky integrators", TKM takes into account the history of the presented vectors in the activation value $U_i(n, d)$ for the neuron i at step t and with d a time constant. The activation value is defined as a recursive sum defined by the the equation1 (for further detail see [8]).

$$U_i(t, d) = dU_i(t-1, d) - \frac{1}{2}\|x(t) - w_i(t)\|^2 \tag{1}$$

where $x(t)$ is the input vector a time t and $w_i(t)$ the weight vector on the neuron i at time t. Then the best matching units is determined by the neuron maximal activation value.

An amelioration of this algorithm is the Recurrent Self-Organizing Map (RSOM [9]), it consists of a "generalization" of the leaky integrator to the whole vector, the activation $U_i(.)$ is so replaced by a vector $y_i(.)$ as described in the equation 2, for the remaining details RSOM is similar to TKM [10].

$$y_i(t, \alpha) = (1-\alpha)y_i(t-1, \alpha) + \alpha(x(t) - w_i(t)) \tag{2}$$

With α the analogous value of $(1-d)$ in the TKM. Also using the leaky integrators, the SOM with Temporal Activity Diffusion (SOMTAD [11]), which is based on biological model of activity diffusion. In this algorithm the SOM structure is used to memorize the activity diffusion.

Another algorithm with a different approach is the Recursive SOM (RecSOM [12]). In this algorithm, each output neuron has a weight vector $w_i(t)$ representing its position in the distribution and a context vector $c_i(t)$ containing map activations from the previous time step $y(t)$. $w_i(t)$ is updated as in the SOM algorithm [5], whereas the context vector is updated by the learning rule 3. So the best matching unit is based on a distance measure d depending on the weight vector and the context vector (for more details see [12]).

$$\Delta c_i(t) = \gamma h_{ik}(y(t-1) - c_i(t)), \tag{3}$$

where $h_{ik}(t)$ is the neighborhood function as in the SOM algorithm.

Inspired from the RecSOM, the Merge SOM (MSOM [13]). The principal difference consists in the context vector. It represents here a recursive sum between the weight vector and the context vector from the previous iteration initially set to zero (Eq. 4).

$$c(t) = \gamma c(I(t-1)) + (1-\gamma)w(I(t-1)) \tag{4}$$

with $I(t)$ the best matching unit at time t and γ a time constant (for more details see [13]).

All the presented methods and many others (Neural Gaz [14], SOM for Structured Data SOMSD [15], Hyperbolic SOM [16]...) have in common that they are inspired from the SOM algorithm initially proposed by Kohonen in the 1980s. As we said, the quantization is just the first step, after this step a predictor should be used. In the beginning, the idea were to use a very simple models such as the AR predictors since the time series supposed to be locally stationary [3]. Then with more complex and chaotic data, these models were replaced by Neural Networks [6] (such as MLP). However, the RNNs was not introduced with this kind of approach despite of the good results obtained by these n. In neural networks as predictors next section we will discuss how RNN's could be used after a the Vector Quantization phase.

3 RNN with Local Approach

The first (globally) RNN was proposed by Jordan [17], in which we have four type of layer : Input layer, hidden layer, output layer and context layer. Globally, the network is like a MLP one, except the context layer which contains also a loop back in each neuron (see Fig. 1. This networks was improved by Elman [17] by adding a feedback connection between the hidden layer and the context layer. It's possible also to use a feedback connection between output layer and the context layer 2. Although, many other types of RNN's are possible such as a fully connected neural

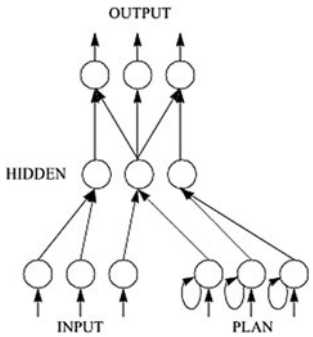


Fig. 1 Jordan recurrent neural network

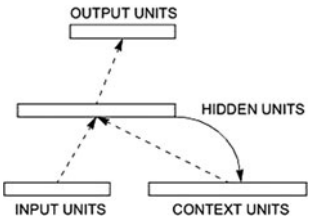


Fig. 2 Elman recurrent neural network

network [6], a fully connected hidden layer[] or the Long-Short Term Memory networks (LSTM [18]).

In feed forward Neural Networks (i.e MLP) the information is forwarded from layer to layer with the Back Propagation(BP [2]) algorithm. However the prediction value depends only on the k past values of the temporal window. However, with Recurrent Neural Networks the information feedback is allowed. Which means that in addition to the input data coming from the temporal window, an other type of information is a combination of the old values coming from the internal loop in the network. This can be one of the reasons that RNN' have demonstrated a good performances and special robustness face to long term dependencies.

In consequence the BP algorithm has to include the temporal notion, explained by the information propagation into a simple connection between two neurons and which should take one time step. One algorithm used for RNN's is the Back Propagation Trough Time (BPTT [2]). The BPTT is based on a gradient descent of the error for the weight modifications $\Delta w_{ij}(t_1, t_l - 1)$ described in the Eq. (5):

$$\Delta w_{ij}(t_1, t_l - 1) = -\eta \sum_{\tau=t_1}^{t_l-1} \frac{\partial E(t_1, t_l)}{\partial net_i(\tau)} s_j(\tau) \tag{5}$$

The gradient error $\dfrac{\partial E(t_1, t_l)}{\partial net_i(\tau)}$ is calculated depending on value of τ:

- for $\tau = t_l - 1$

$$\frac{\partial E(t_1, t_l)}{\partial net_i(\tau)} = \begin{cases} (s_i(t) - d_i(t)) f_i^{'}(net_i(\tau)) & \text{if } i \text{ is an output neuron} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

- for $t_1 \leq \tau \leq t_l - 1$

$$\frac{\partial E(t_1, t_l)}{\partial net_i(\tau)} = \begin{cases} (s_i(t) - d_i(t) + \sum_{j \in succ(i)} \frac{\partial E(t_1, t_l)}{\partial net_j(\tau+1)} w_{ij}(\tau + 1) f_i^{'}(net_i(\tau)) \\ \text{if } i \text{ is an output neuron} \\ \sum_{j \in succ(i)} \frac{\partial E(t_1, t_l)}{\partial net_j(\tau+1)} w_{ij}(\tau + 1) f_i^{'}(net_i(\tau)) \\ \text{otherwise,} \end{cases} \tag{7}$$

The internal feedback in the network requires that the data should be presented in the learning algorithm sequentially and with a constant time step between entries. However after the Vector Quantization phase the temporal information between vectors is lost as the time step is no longer be constant for each RNN associated to a specific cluster. In order to keep a constant time step, we propose to present all the vectors to each RNN and to add a supplementary information $\delta_{i,k}$ defined by:

$$\delta_{i,k} = \begin{cases} 1 & \text{if vector } i \in \text{cluster} k \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

For each RNN of a cluster i, all the data of the learning step will be presented. So the learning rule (Eqs. 6 and 7) of BPTT algorithm should be updated in order to take into account only the vectors belonging to their appropriate clusters. The $\delta_{i,k}$ is added in Eqs. (9) and (10) in order to indicate when a gradient descent should be calculated for a vector.

- for $\tau = t_l - 1$

$$\frac{\partial E(t_1, t_l)}{\partial net_i(\tau)} = \begin{cases} (s_i(t) - d_i(t))\delta_{i,k} f_i^{'}(net_i(\tau)) & \text{if } i \text{ is an output neuron} \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

- for $t_1 \leq \tau \leq t_l - 1$

$$\frac{\partial E(t_1, t_l)}{\partial net_i(\tau)} = \begin{cases} \left[(s_i(t) - d_i(t)) + \sum_{j \in succ(i)} \frac{\partial E(t_1, t_l)}{\partial net_j(\tau+1)} w_{ij}(\tau + 1)\right] \delta_{i,k} f_i^{'}(net_i(\tau)) \\ \text{if } i \text{ is an output neuron} \\ \sum_{j \in succ(i)} \frac{\partial E(t_1, t_l)}{\partial net_j(\tau+1)} w_{ij}(\tau + 1)\delta_{i,k} f_i^{'}(net_i(\tau)) \\ \text{otherwise,} \end{cases} \tag{10}$$

This modification in Eqs. (9) and (10) can be interpreted by a double dependence of the target function. If we consider $\hat{g}(.)$ the function to approximate in specific to a cluster of data, this function depends not only on the vector $X(t)$ belonging to this cluster, but also on the other data presented to neural network in the previous steps. Let $\hat{x}(t)$ be the predicted value, we can say 11:

$$\hat{x}(t) = \hat{g}(X(t), \widetilde{h}(t)), \tag{11}$$

4 Experimental Results

In this section we have applied the algorithm proposed in two different times series. The first experiments are with the Mackey-Glass Times series (MG). The second experiments are with the Laser data.

The MG chaotic time series can be generated from the MG delay-differential Eq. (12) [19]:

$$\dot{x}(t) = \frac{\alpha x(t-\tau)}{1+x^c(t-\tau)} - \beta x(t), \tag{12}$$

the benchmark sets can be generated using a four-point Runge-Kutta method and with initial condition $x(t) = x_0\, for\, t < 0$, a, b and c are equation parameters. For MG-17 the τ is set to 17. The Laser data is a set from the Santa Fe time series prediction competition [20]. It consists of one-dimensional data recorded from a far-Infrared (FIR) laser in a chaotic state.

For the evaluation of the forecasting techniques, we have used the mean square error (MSE). It is a quadratic error function that gives more weight to large error. This can be defined by the Eq. (13).

$$MSE = \frac{\sum_n^{i=1} (s(i) - d(i))^2}{n}, \tag{13}$$

with $s(i)$ and $d(i)$ respectively the predicted and desired value for the input i.

Table 1 compares the performance of the proposed algorithm with three well known techniques. The first one is an MLP neural networks with five hidden neurons, methods predicting a value by observing T past values (temporal window). The same MLP neural networks is used with a SOM algorithm, composed of four unit (2×2) in the output grid. Then a RNN with a fully hidden layer of eight neurons and with a temporal windows of a T past values is used . This same RNN is also tested with a SOM algorithm defined also by (2x2) units in the output grid.

The input pattern of the SOM is defined as:

Table 1 Prediction error

Method	Laser MSE	MG-17 MSE
MLP (T = 4)	$1.65 \cdot 10^{-1}$	$7.15 \cdot 10^{-2}$
MLP (T = 6)	$2.68 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$
RNN (T = 4)	$3.18 \cdot 10^{-1}$	$5.89 \cdot 10^{-2}$
RNN (T = 6)	$1.99 \cdot 10^{-1}$	1.07
SOM+MLP (T = 4)	$3.45 \cdot 10^{-2}$	$3.94 \cdot 10^{-2}$
SOM+MLP (T = 6)	$3.22 \cdot 10^{-2}$	$5.32 \cdot 10^{-2}$
SOM+RNN (T = 4)	$1.67 \cdot 10^{-1}$	$2.36 \cdot 10^{-2}$
SOM+RNN (T = 6)	$1.39 \cdot 10^{-1}$	$4.94 \cdot 10^{-1}$

$$x(t) = (x(t), x(t-1), ..., x(t-\tau)), \tag{14}$$

with τ the value of the temporal window. The same vector is then passed to the recurrent neural network with the value of $\delta_{i,k}$. The table 1 shows that a Vector Quantization improve considerably the prediction error, first the MLP neural networks all ready proved in the literature, but also for the Recurrent Neural Networks. The experiments shows that despite the varying time step, the RNN can be improved with the SOM algorithm, also true when the RNN outperforms the MLP.

5 Conclusion

In this paper we have proposed to use a RNN with a vector quantization (SOM in our case). Then this algorithm was tested on two type on data (Laser and MG-17). The proposed methods was compared to well known techniques.

The empirical results suggests that the RNN can perform a good results. This is encouraging to employ more complex RNNs used by the past [6] that have showed a very hight performance. Future research will focus on the prediction on specific clustering algorithm to be applied with RNNs.

References

1. Reinsel, G.C., Box, G.E.P., Jenkins, G.M.: Time Series Analysis: Forecasting and Control. Englewood Cliffs, NJ: Prentice Hall (1994).
2. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. Parallel distributed processing: explorations in the microstructure of cognition, vol. 1. Foundations (1986) 318–362.
3. Walter, J., Riter, H., Schulten, K.: Nonlinear prediction with selforganizing maps. In International Joint Conference on Neural Networks (IJCNN), vol. 1. (1990) 589–594.
4. Gray, R.M., Neuhoff, D.L.: Quantization. IEEE Trans. Inf. Theory **44**(6) (1998) 2325–2383.
5. Kohonen, T.: Self-organization and associative memory: 3rd edition. New York, NY: Springer-Verlag New York, Inc. (1989).
6. Boné, R.: Recurrent neural networks for time series forecasting. PhD thesis, Université de Tours, Tours, FRANCE (2000).
7. Roman Jesus Martin-Merino Manuel. A new som algorithm for electricity load forecasting. In International Conference on Neural Information Processing. (2006) 995–1003.
8. Chappell, G.J., Taylor, J.G.: The temporal kohonen map. Neural Netw. **6**(3) (1993) 441–445.
9. Varsta, M., Millán, J.D.R., Heikkonen, J.: A recurrent self-organizing map for temporal sequence processing. In ICANN '97: Proceedings of the 7th International Conference on Artificial Neural Networks. London: Springer-Verlag (1997) 421–426.
10. Varsta, M., Heikkonen, J., Lampinen, J., Millán, J.D.R.: Temporal kohonen map and the recurrent self-organizing map: Analytical and experimental comparison. Neural Process. Lett. **13**(3) (2001) 237–251.
11. Euliano, N.R., Principe, J.C.: A spatio-temporal memory based on SOMs with activity diffusion. In Oja, E., Kaski, S., eds.: Kohonen Maps. Amsterdam: Elsevier (1999) 253–266.
12. Voegtlin, T.: Recursive self-organizing maps. Neural Netw. **15**(8–9) (2002) 979–991.
13. Strickert, M., Hammer, B.: Merge som for temporal data. Neurocomputing. **64** (2005) 39–71, 2005. Trends in Neurocomputing: 12th European Symposium on Artificial Neural Networks 2004.

14. Martinetz, T.M., Berkovich, S.G., Schulten, K.J.: 'neural-gas' network for vector quantization and its application to time-series prediction. IEEE Trans. Neural Netw. **4**(4) (1993) 558–569.
15. Sperduti, A.: Neural networks for adaptive processing of structured data. In ICANN '01: Proceedings of the International Conference on Artificial Neural Networks. London: Springer-Verlag (2001) 5–12.
16. Ritter, H.: Self-organizing maps on non-euclidean spaces. In Oja, E., Kaski, S., eds.: Kohonen Maps. Amsterdam: Elsevier (1999) 97–108.
17. Elman, J.L.: Finding structure in time. Cognitive Science **14**(2) (1990) 179–211.
18. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8) (1997) 1735–1780.
19. Mackey, M.C., Glass, L.: Oscillation and chaos in physiological control systems. Science **197**(4300) (July 1977) 287–289.
20. Weigend, A.S., Nix, D.A.: Prediction with confidence intervals (local error bars). Proceedings of International Conference on Neural Information Processing (ICONIP'94) (1994) 847–852.

Single-Trial Electrical Perception Recognition via Single-Channel Electroencephalogram

Yi Qiu, Guiping Dai, Yuping Miao, Ruifen Hu, and Guang Li

Abstract Single-trial classification of electroencephalogram (EEG) has received increasing attention for psychological and engineering research, particular the brain computer interface (BCI). A novel method to recognize the EEG pattern induced by an electrical current stimulus was studied. Experimental results showed the electrical stimulating pattern can be recognized via single-trial and single-channel EEG using artificial neural networks. For a BP neural network, the correction ratio was better than 90%. The method has potential to use for the objective study on human perception.

Keywords EEG · Single-trial · Artificial neural network · Perception

1 Introduction

The study of the relationship between EEG patterns and stimuli is promising for understanding the mechanism of neural information processing, and has potential for the human sensory function estimation and BCI applications. Single-trial and single-channel EEG pattern recognition is a challenge for signal processing and machine learning [1]. Based on well known event related potentials (ERPs) [2], we developed a method to classify electrical perception pattern via single-trial single-channel EEG. This method can assess and quantify sensory function in patients with neurologic symptoms or in those at risk of developing neurologic disease.

Y. Qiu (✉)
Department of Control Science and Engineering, National Key Lab of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China
e-mail: endlessrain@zju.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_93,

2 Methods

The procedure to recognize the EEG pattern evoked by electrical current is illustrated in Fig. 1. According to spectrum analysis, principal energy of ERPs evoked by electrical stimuli is concentrated on the frequency band of 0–3 Hz. So the single-channel EEG was firstly filtered with a low-pass filter of 0–3 Hz, and then segmented properly. The Euclidean distances between the segmented EEG and the EEG templets with and without stimulus, which are indicated as X_1 and X_2 respectively, were calculated to form a feature vector of the single-trial and single-channel EEG signal. Classified by a BP neural network, the EEG with or without electrical perception can be discriminated.

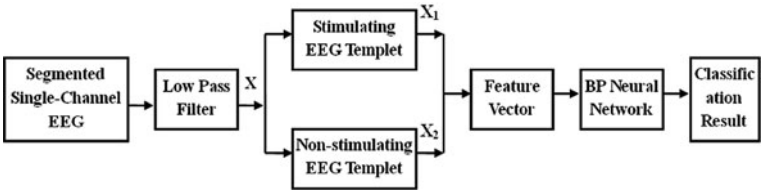


Fig. 1 The procedure of EEG signal processing for electrical perception recognition

2.1 Data Acquisition

Three male students, age of 20, 21 and 26, volunteered as subjects. All subjects were free of neurological or psychiatric disorders. The subject seated on a comfortable chair with his eyes closing and arms resting on a table during the experiments. After the current perception threshold (CPT) of the subject was estimated [3], a stimulating electrode of 10 mm diameter was applied at the index finger's end joint of the subject's left hand to provide an electrical current stimulus. 250 Hz sine wave lasting 2 s and about 3 times as the subject's CPT was applied on the finger of the subject, while the CPz channel EEG was recorded with the reference linked mastoids and a forehead ground. The EEG amplifier used was NuAmps Digital Amplifier (Model 7181) purchased from the Neuroscan Compumedics Limited, Texas, USA. The sampling rate is 500 Hz. A 0.5–70 Hz pass-band filter was applied. All impedances were kept below 5 KΩ during EEG recording and the time when the electrical stimulus started was recorded synchronously as time reference.

2.2 Templet Set-Up and EEG Segment

It is difficult to distinguish whether ERP existing within original EEG as shown in Fig. 2a, c. After 20-times average of EEG without stimulating and with stimulating, the average filtered EEGs were obtained and to be used as the non-stimulating

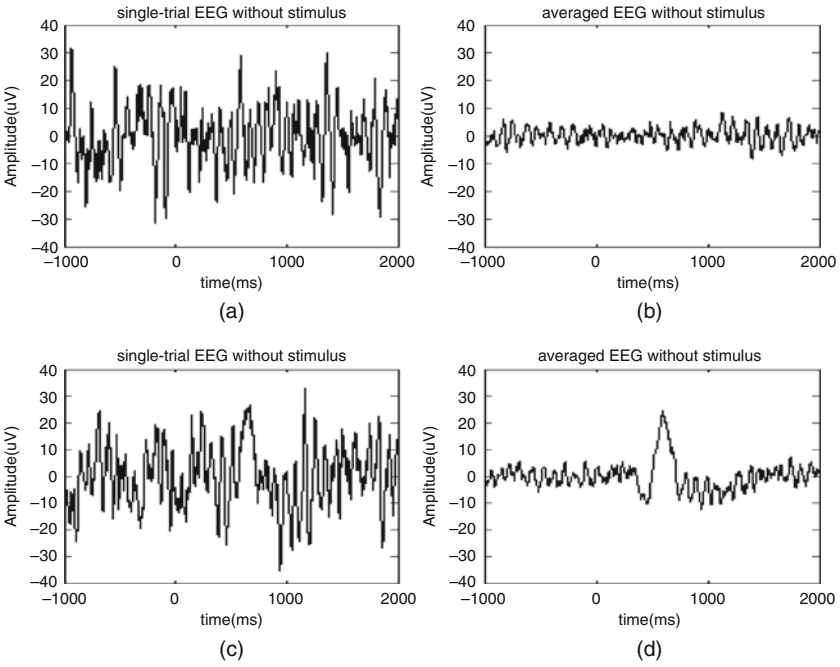


Fig. 2 EEG waveforms, where 0 ms on x-axis indicates the time when stimulus started for (**c**) and (**d**). (**a**) Original EEG without stimulus, (**b**) Averaged EEG without stimulus, (**c**) Original EEG with stimulus, (**b**) Averaged EEG with stimulus (ERP)

(Fig. 2b) and stimulating (Fig. 2d) templets, respectively. After 20-times average, the electrical stimulus induced ERP can be clearly seen about 600 ms after the stimulus started (shown as in Fig. 2d). 128 samples of EEG signal around the peak of ERP were segmented for further analysis.

2.3 Feature Extraction

For each subject, 20 times EEG with stimuli were recorded and averaged as the stimulating EEG templet, while 20 segments of EEG without stimulus were averaged as the non-stimulating EEG templet. These EEG segments were also utilized later to form training data set for pattern classification.

For each EEG segment to be estimated, the means of the Euclidean distances to stimulating and non-stimulating EEG templets were respectively calculated to construct a 2-dimensional feature vector. So the dimension of the EEG segment consisting of 128 points was reduced. The 2-dimensional project of the feature vectors as shown in Fig. 3.

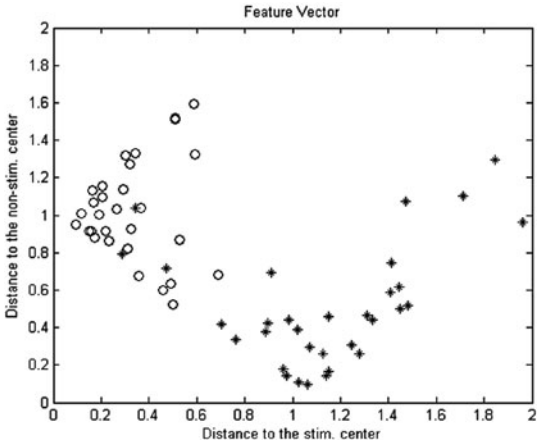


Fig. 3 Distribution of feature vectors (*star* – stimulating EEG, *circle* – non-stimulating EEG)

2.4 Single-Trial Estimation Based on BP Neural Network

The back-propagation algorithm with adaptive learning parameters for feed-forward neural network learning [4], which consists of two neurons in the input layer, one neuron in the output layer and six neurons in the single hidden layer, was used to classify the EEG segments with or without electrical perception. For each subject, 40 EEG segments used to form the templets (20 for each EEG templet) were used as the training set. EEG segments obtained from different trials were used for test. For all these subjects, the correction ratios of classification are 90.0, 95.8 and 93.3%, respectively.

3 Conclusion

Efficient features of single-trial electrical perception can be extracted based on single-channel EEG using subject's ERP obtained in advance as the templet. Utilizing BP neural network as classifier, the correction ratio better than 90% can be achieved. This method can be utilized to judge whether an electrical stimulus is perceived, so that the current perception threshold can be tested objective. Furthermore, this method is promising for understanding the mechanism of neural information processing of perception.

Acknowledgments This research is supported by the Natural Science Foundation of China (Grants No. 60874098 and 60911130129), the National High Technology Research and Development Program of China (863 Program 2007AA042103) and the National Creative Research Groups Science Foundation of China (NCRGSFC: 60421002).

References

1. Guan, J., Chen, Y., Lin, J.: Single-trial estimation of imitating-natural-reading evoked potentials in single-channel. Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. (2005) 2052–2055.
2. Ng, K.S., Yang, H., Kim, S.: Hidden pattern discovery on event related potential EEG signals. BioSystems **97**(2009) 15–27.
3. Shy, M.E., Frohman, E.M., So, Y.T.: Quantitative sensory testing. Am. Acad. Neurol. **60**(2003) 898–904.
4. Kathirvalavakumar, T., JeyaseeliSubavathi, S.: Neighborhood based modified backpropagation algorithm using adaptive learning parameters for training feedforward neural networks. Neurocomputing (2009) in press.

Sleep Quality Analysis Based on HHT

Lengshi Dai, Guiping Dai, Wuming Zhang, Haigang Zhu, and Guang Li

Abstract This paper introduces a method to estimate sleep quality via EEG analysis based on HHT. Sleep stages were classified throughout the method where instantaneous frequencies of sleep EEG with physical meaning and energy frequency distribution are used as feature parameters for each stage computed with the HHT. Meanwhile, the well known psychological test called PSQI was carried out for comparison. The experimental results showed that the EEG analysis based on HHT is promising for objective evaluation of sleep quality comparable to PSQI.

Keywords Sleep stage · Hilbert-Huang transform · PSQI · EEG · EOG

1 Introduction

Although the international standard entitled with R&K Rules that classifies sleep stage based on electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) analysis has been established since 1968. Together with the thriving of computer science and IT industry, many methods to extract the features of the electrophysiological signals leading to final judgment of sleep stages including frequency analysis method, data complexity, entropy analysis etc. have been studied. In this paper, the Hilbert-Huang transform (HHT) was utilized to analyze EEG signal for sleep quality estimation. Introducting the `intrinsic mode functions′ based on local properties of the signal, it makes the instantaneous frequency meaningful to eliminate the need for spurious harmonics to represent nonlinear and non-stationary signals for complicated data sets [1].

According to the R&K rules, which is a world-wide criterion for sleep stages classification published in 1968, human sleep can be divided into six phases: waking

L. Dai (✉)
Department of Control Science and Engineering, National Key Lab of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China
e-mail: all-stardai@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_94,

(Stage W), the rapid eye movement (REM) stage and four non-REM stages (from the lightest Stage 1 to Stage 4. The scoring of sleep quality, usually accomplished by well-trained personnel, consists in classifying all 30 s pieces of a sleep recording into one of six stages [2]. The stage is defined according to the EEG waveform and eye movement [3]. In order to identify REM from non-REM efficiently, the angular departure from the baseline in EEG is selected as an extra criterion [4]. Each stage has its own distinctive features in need of obtaining from EEG, EOG and EMG, for example, stage I abounds with larger amount of Alpha wave, the feature of Stage II is the appearance of spindle wave, stage III and stage IV are mainly defined as delta wave with relatively large amplitude. Therefore, the key to the sleep stage classification is to choose a proper algorithm that is able to collect the specific energy composition of instantaneous frequency.

PSQI is a well known psychological test designed by Pittsburgh University, USA, for the sake of measuring the sleep quality in the recent 1 month. This currently used index includes 19 questions towards oneself and five ones answered by others (among these questions, only eighteen of those will be taken into use for scoring), covering all sleep respects, such as “time to fall into asleep”, “times wake up at night” and so on.

Overall, the index is made up of seven ingredients divided to be A to G alphabetically, where each ingredient is supported by one or more questions in the raw index answers replied objectively by the participator. While the final score is the cord to ultimately answer how the sleep quality is, this score has reverse ratio to the real sleep quality: that means the higher the score is, the worse the sleep quality is.

2 Experiment and Results

Nine volunteers, 5 males and 4 females, age from 28 to 68, were recruited as subjects. The EEG recording was executed at night during the sleep duration from the regular time for bed till the sleeper’s natural wake up.

In order to test the efficiency of the method, the mattresses specially designed for improving sleep quality were used to try to make sleep quality different for same person. The mattresses were provided to the subjects free of charge for 60 days. No sleeping pill was used. To check the veracity of the HHT based EEG analysis method, the subject was interviewed to complete the PSQI for comparison.

The EEG amplifier used was NuAmps Digital Amplifier (Model 7181) purchased from the Neuroscan Compumedics Limited, Texas, USA. The sampling rate was 250 Hz. A 0.5–30 Hz band pass filter was applied for the EEG channel, and 40 Hz low pass filter for EOG. Due to the fact that the noise mainly exists in the delta band where the amplitude but not the frequency is the dominant factor for the discrimination of sleeping stages, there is no need for approaches to eliminate the noise other than the band pass filter. The artificial rejection was done by the Neuroscan software.

The Hilbert–Huang transform (HHT) is utilized to analyze the EEG recordings. HHT is an algorithm to decompose a signal into so-called intrinsic mode functions

(IMF), and obtain instantaneous frequency data [1, 5]. It is especially effective to describe nonlinear and non-stationary distorted waves in details. In contrast to conventional Fourier transform, HHT performs better when the signal is nonlinear. Further on, for users, the problem that the resolution for time and frequency could not be achieved simultaneously which had greatly weakened the effect of FFT would no longer happen to HHT. It means we are able to insure the resolution in time scale without sacrifice the precision of frequency.

The EEG was firstly divided into epochs of 30 s which were then classified into one of sleep stages [6]. For each segment, following features were extracted to staging the epochs (the total power equals to 0–30 Hz):

(1) Amplitude: the average of the absolute value of the signal and the peak to peak value for each second.
(2) Spindles: presence is assessed using the ratio of power in the sigma band (11.5–15 Hz) and the total power, relative to a similar ratio of background EEG [6].
(3) Alpha: the ratio of power in the alpha band (8.0–11 Hz) to the total power.
(4) Delta: the ratio of power in the alpha band (0.5–4 Hz) to the total power.
(5) The REM activity: During REMs the EOG signal shows a steep rising and a flat falling edge with its peak surpass 40 μV and the slope with a 45°or greater departure from the base line.

Thus, the final image, which representing the entire process of sleep, was completed after the combination of consecutive segments labeled with sleep stages.

In order to describe sleep efficiency, the indices including the ratio of time asleep to the whole-night sleep duration (T), deep sleep ($\frac{Stage\ III+stage\ IV}{T} \times 100\%$), the sleep latency period which indicating the time for a subject falling into sleep, sleep interruption were calculated. The results are given in Table 1. And some main indices of one subject during 2-month are shown in Fig. 1.

Table 1 The sleep quality for volunteers before (B) and after (A) using the mattress, where Stages III and IV were combined to collaborate to represent the deep sleep

			Sleep efficiency (%)		Deep sleep (%)		Sleep latency (min)		Sleep interruption time(s)		PSQI (points)	
No.	Gender	Age	B	A	B	A	B	A	B	A	B	A
1	M	42	89.0	91.4	14.7	16.6	14.7	14.3	3.0	1.3	8	3
2	F	40	94.7	94.5	15.6	21.8	18.7	13.5	2.3	1.0	7	4
3	M	42	86.9	89.7	15.4	21.2	23.0	13.2	2.0	0.7	7	4
4	F	37	89.3	90.6	14.4	15.3	23.8	14.2	2.0	1.7	5	5
5	F	29	81.3	91.3	22.7	22.6	10.8	10.5	1.7	1.0	4	2
6	M	32	97.8	96.2	22.2	22.7	10.3	17.3	1.3	0.3	2	2
7	M	27	86.8	89.6	18.5	25.2	14.5	15.2	2.0	1.0	6	3
8	M	68	85.9	93.7	2.7	4.6	37.8	12.3	5.0	1.3	15	8
9	F	28	88.0	92.1	21.0	22.9	13.8	18.0	3.7	1.0	9	3

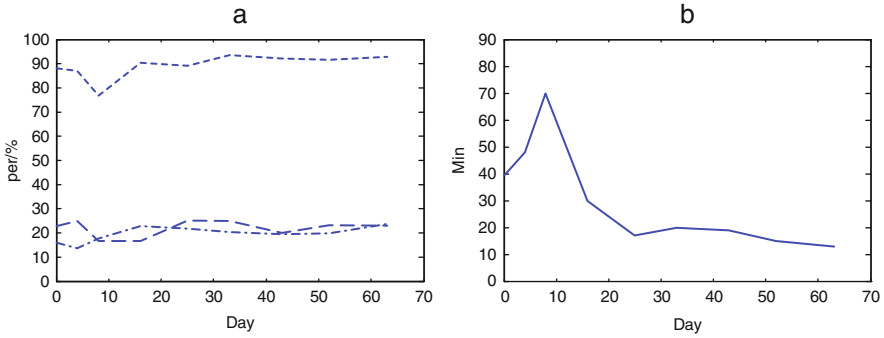


Fig. 1 Track for a subject volunteer (**a**) sleep efficiency (*dotted line*); deep sleep (*dash dotted line*); REM sleep (*dashed line*) (**b**) sleep latency

3 Conclusion and Discussions

The indices relating to sleep quality, which including deep sleep percentage, sleep latency and sleep interruption time, were calculated mainly according to the HHT results of sleeping. Based these indices, the sleep efficiency was estimated quantitatively. It can be seen from Table 1 that for same subject's the sleep quality including the sleep efficiency, deep sleep percentage, sleep latency time and sleep interruption times changed with PSQI. It is clear from Table 1 that the sleep efficiency does not exactly match the PSQI though they are approximately coincident with each other. At same sleep efficiency level, more deep sleep and less sleep latency means better sleep quality. In order to evaluate sleep quality more accurate, a new sleep index which combining all those parameters mentioned above should be defined.

Acknowledgment This research is supported by the Natural Science Foundation of China (Grants No. 60874098 and 60911130129), the National High Technology Research and Development Program of China (863 Program 2007AA042103) and the National Creative Research Groups Science Foundation of China (NCRGSFC: 60421002).

References

1. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., et al: The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A **454** (1998) 903–995.
2. Šušmáková, K., Krakovská, A.: Classification of waking, sleep onset and deep sleep by single measures. Meas. Sci. Rev. **7**(4, Sec 2) (2007) 34–38.
3. Rechtschaffen, A., Kales, A.: A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Washington, DC: Government Printing Office, Public Health Service (1968).

4. Hori, T., Sugita, Y., Koga, E., et al.: Proposed supplements and amendments to 'a manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', the Rechtschaffen & Kales (1968) standard. Psychiatry Clin. Neurosci. **55** (2001) 305–310.
5. Huang, N.E., et al.: Hilbert-Huang transform. Scholarpedia **3**(7) (2008) 2544.
6. Agarwal, R., Gotman, J.: Computer-assisted sleep staging. IEEE Trans. Biomed. Eng. **48**(12) (2001) 1412–1423.

Modelling the Stroop Effect: Dynamics in Inhibition of Automatic Stimuli Processing

Nooraini Yusoff, André Grüning, and Antony Browne

Abstract In this study, we simulate the dynamics of suppressing an automatic stimulus processing which interferes with a different non-automatic target task. The dynamics can be observed in terms of interference and facilitation effects that influence target response processing time. For this purpose, we use Hopfield neural network with varying attention modulation in a colour-word Stroop stimuli processing paradigm. With the biologically realistic features of the network, our model is able to model the Stroop effect in comparison to the human performance.

Keywords Neurodynamics · Automatic processing · Stroop effect · Hopfield neural network

1 Introduction

Distraction in processing of a target response can occur not only due to properties of irrelevant stimulus, but also through the relationship of the stimulus properties to the target-finding properties that attract the attention and resulting conflict [1]. Some environments can create automatic responses to certain classes of stimuli resulting in distraction from the intended task. This can also lead to priming effects as results of automatic processing of a stimulus (prime) that would distract or facilitate the processing of the later stimulus. The prime stimulus can be a cue if it is congruent with the later stimulus and facilitates its response; otherwise it is a distractor if both are conflicting, resulting in interference in processing. Having such dynamic effect will lead to variability in response processing time. Whenever interference occurs, the unintended automatic response has to be suppressed, using higher attentional control resources in producing the target response to the intended stimulus. Hence,

N. Yusoff (✉)
Department of Computing, Faculty of Engineering and Physical Sciences, University of Surrey, GU2 7XH Surrey, UK
e-mail: n.yusoff@surrey.ac.uk

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_95,

in this study, we attempt to simulate the dynamics of the priming effect through modulation of attention in a colour-word Stroop paradigm [2].

The Stroop paradigm demonstrates human cognitive processing when one process dominates and inhibits the response of another process. Such environment involves automatic processing of a stimulus, distracting or sometimes facilitating the response to the intended stimulus. For example, in a colour-word Stroop test, the stimuli are the coloured colour-words in three conditions; control (e.g. a word RED written in black or a non colour-word – e.g. CAT written in red), conflicting (e.g. a word RED written in green) and congruent (e.g. a word RED written in red). Subjects are asked either to read the colour-words or to name their colours while the reaction time of performing the task is observed.

The results from Stroop studies show increased reaction time in naming the colour of the printed colour-word in the conflicting condition, while the subject could easily read the word without any significant distraction by the colour it is written in (e.g. [2–4]). The conjectured explanation is that the asymmetric processing is due to automaticity of the word stimulus that is always processed first prior to the intended task stimulus, the colour. This consequently leads to the priming effects, whereby to name the colour of the word, one needs to inhibit the prime processed stimulus (the colour-word). The conflicting stimulus will increase the response processing time in colour naming; on the other hand the congruence of both will facilitate the later stimulus reducing the response processing time. For simulating the interference of the stimuli, we use a Hopfield neural network (HNN) with asynchronous pattern update procedures [5]. The Hopfield network is chosen for several reasons; we address the Stroop phenomenon as an association problem, the competition and cooperation of Stroop stimuli meets the pattern processing nature of the Hopfield network and the recall algorithm in Hopfield is biologically realistic.

2 Methods

To simulate the dynamics in inhibition of automated responses, we created a HNN with a memory of four ($M = 4$) random 56-bit ($n = 56$) patterns. Each pattern represents an association of the colour concept consisting of two congruent components, the colour-word (<WORD>, 16-bit) and the visual colour (<colour>, 16-bit) (based on findings by Folk et al. [1]), and the attentional resource (<Attention>, 24-bit). "<Attention>" models how much a subject attends to a task. If the subject is less attentive less on-bits are inherited from the memory pattern to the test pattern. With such association, we postulate that, each colour concept has some degree of attention resource, cooperating among them simultaneously in recalling a target pattern.

There are two main phases involved in our HNN based Stroop simulation; training phase and testing phase (based on mathematical computation as in [5, 6]). During the training phase, all memory patterns (a pattern noted as x) are correlated to each other using the equation in (1). The correlation derives a set of weights ($W_{56\times56}$) as a product of pattern vector associations.

$$w_{ij} = \sum_{k=1}^{M} x_i(k)x_j(k),\ i \neq j,\ w_{ii} = 0,\ i,j = \overline{1,n}. \tag{1}$$

Once a set of association weights is obtained, in the testing phase, a test pattern, *xr*, is presented to the system. In our model, *xr*s represent the Stroop stimuli with an amount of attention depending on the task to recall the target colour concept. For Stroop stimuli representations, from each memory set, 20 test patterns are generated to observe the recall performance of a Stroop task (word reading or colour naming). The performances are observed under three conditions of stimuli; control – absence of irrelevant stimulus to the attended task (e.g. for a word reading; <Attention><RED><minimal noise>, 4 test patterns), conflicting – incongruent colour concept (e.g. <Attention><RED><green>, 12 test patterns) and congruent – compatible colour concept (e.g. <Attention><RED><red>, 4 test patterns).

For experimental setup, the initial activation ("on" bits) of a pattern depends on the task; word reading (WR) or colour naming (CN) is assumed to be signalled by another external system. For activation of bits in any components of a pattern, it refers to the percentage of similarity of "on" bits in the correspond memory. For instance, let a <WORD> component in a test pattern inherits *n*% random activations from its memory with 12 bits "on" (out of 16 bits), then there would be $n\% \times 12$ "on" bits in the component. For a WR task, the <WORD> component (relevant stimulus) inherits 75% random activation of its memory, while having random activation of the <colour> component (irrelevant stimulus) (ranging from 0 to 25% on-bits, with random uniform distribution). In contrast, for a CN task, the <colour> stimulus inherits 25% random activation from its memory, with random activation of <WORD> component (irrelevant stimulus) (ranging from 0 to 75% on-bits, with random uniform distribution). Greater maximum initial activation of word stimulus in CN is to simulate its automaticity that we predict would cause interference in processing the response to the task. For this reason, the colour stimulus can only survive with higher attention i.e. a test <Attention> that is more similar to the correspond <Attention> of a memory pattern. The initial activation applies to all conditions except the control stimuli with maximum of 25% noise in irrelevant stimulus.

In contrast with some other Stroop models (e.g. [3, 4]) emphasizing the influence of neuron connection pathways (weights) in response processing, ours considers the influence of attention. In our model, the modulation of attention in a recall process is simulated through a part-pattern (<Attention> bits) completion. For our simulations, we varied the initial activation of attention from 0% (all off-bits) to 100% (identical attention component from correspond memory pattern). The dynamics in HNN asynchronous update would eventually complete the initial random activation of <Attention> vector corresponding to a target memory.

To recall a target memory pattern, at any given time, each bit (representing a neuron in biological system) in *xr*, receives the total activation (*net*, computed using (2)) from others through the HNN asynchronous update mechanism:

$$net_i(t) = \sum_{i \neq j} w_{ji}(xr_j(t)), \tag{2}$$

where $net_i(t)$ is the *net* input to neuron i at time t, w_{ji} is the connection strength between neuron j to neuron i, xr_j is the state of neuron j (+1 or −1). In an update cycle, the state of neuron i is readjusted according to 3.

$$xr_i(t+1) = \Theta\left(net_i(t)\right) = \{+1,\ net_i(t) > 0;\ -1,\ net_i(t) < 0;\ xr_i(t)\ , \\ net_i(t) = 0\}\ . \tag{3}$$

The number of bit updates in a cycle with maximum of 300 iterations for a stimulus recalling the closest (measured by Euclidean Distance) target memory is recorded as the reaction time. This simulates the reaction time taken by a subject to perform any of the tasks. Graphs and numbers below are the averages for repetitions of the experiment with 10 different choice of memory sets, and 11 different levels of attention. Asynchronous updates of bits in <Attention>, <WORD> and <colour> vectors simultaneously, simulate the dynamics of cognitive process in Stroop phenomenon for active inhibition, facilitation and interference.

3 Results

As shown in Fig. 1, using the Stroop stimuli to recall their target memories through active inhibition of Hopfield's algorithm, our model predicts the asymmetric stimuli processing in the colour-word Stroop phenomena in comparison to the human performance.

For our model, we consider the average of the reaction time (RT) in processing response for both WR and CN at all levels of initial attention activation ranging from 0 to 100%. After running 110 different simulations (10 memory sets for 11 levels of attention), the model shows that the words are always read faster than colours are named with no significant difference of RT (*ANOVA*: $p > 0.05$) on the stimuli conditions ($\mathrm{WR}_{(RT,\ control)} = 7.05$, $\mathrm{WR}_{(RT,\ conflict)} = 7.26$, $\mathrm{WR}_{(RT,\ congruent)} = 6.53$). Meanwhile there is a significant difference (*ANOVA*: $p < 0.01$) found in CN in all

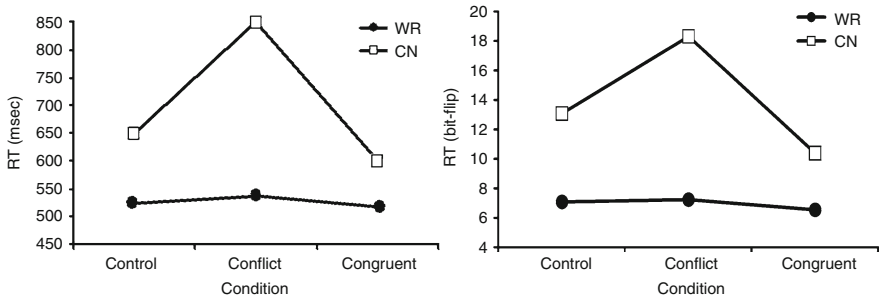


Fig. 1 Performance results for Stroop task averaged over all levels of attention; word reading (*WR*) and colour-naming (*CN*). (*Left*: Results from empirical study after Dunbar and MacLeod [7], *Right*: Results of the model's simulation)

conditions ($WCN_{(RT,\ control)} = 13.05$, $CN_{(RT,\ conflict)} = 18.32$, $CN_{(RT,\ congruent)} = 10.39$), except between the control and the congruent stimuli (*ANOVA* : $p_{(control,\ congruent)} > 0.05$). The interference is obviously observed for CN with increasing time in the conflicting condition whilst the slight benefit of automatic word processing is shown with the decreasing reaction time for congruent condition. In addition to the RT, the frequency (freq) of correct recalls to target responses was also observed. As what we predicted, the same simulations provide consistent results in frequencies of correct recalls with RT, indicating longer processing time leads to higher recall error rate. The correct recalls recorded in WR are as follows; $WR_{freq(control,\ correct)} = 87.82$, $WR_{freq(conflict,\ correct)} = 87.43$, $WR_{freq(congruent,\ correct)} = 91.87$, whilst for CN we obtained; $CN_{freq(control,\ correct)} = 70.64$, $CN_{freq(conflict,\ correct)} = 23.25$, $CN_{freq(congruent,\ correct)} = 81.02$. However in this study, without also neglecting the recall frequencies results, we only focus on the RT as the determinant of the Stroop effect. The results conclude that high inhibition has occurred in CN especially in the conflicting condition due to the incompatibility of the prime stimulus (the word) with the intended stimulus (the colour) having caused the interference, where this has been demonstrated at any level of pre-selective attention in our simulations. On the other hand, a prime stimulus compatible with the target response of the intended task speeds up the processing time.

4 Conclusions

In our approach, we have used a HNN to demonstrate the dynamics of inhibition by an automatic response over an intended stimulus response. The dynamics can be seen as the influence of priming effects resulting from automatic processing. The results showed that our model is able to model the reaction times in the colour-word Stroop paradigm.

References

1. Folk, C.L., Remington, R.W., Johnston, J.C.: Involuntary covert orienting is contingent on attentional control settings. J. Exp. Psychol. Hum. Percept. Perform. **18** (1992) 1030–1044.
2. Stroop, J.R.: Studies of interference in serial verbal reactions. J. Exp. Psychol. **18** (1935) 643–662.
3. Cohen, J.D., Dunbar, K., McClelland, J.L.: On the control of automatic processes: A parallel distributed account of the stroop effect. Psychol. Rev. **97**(3) (1990) 332–361.
4. Phaf, R.H., van der Heijden, A.H.C., Hudson, P.T.W.: SLAM: A connectionist model for attention in visual selection tasks. Cogn. Psychol. **22** (1990) 273–341.
5. Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA **9**(2554) (1982).
6. Popoviciu, N., Boncut, M.: On the Hopfield algorithm. Foundations and examples. General Math. **13**(2) (2005) 35–50.
7. Dunbar, K., MacLeod, C.M.: A horse race of a different color: Stroop interference patterns with transformed words. J. Exp. Psychol. Hum. Percept. Perform. **10** (1984) 622–639.

A New Support Vector Machine Algorithm with Scalable Penalty Coefficients for Training Samples

Jian Zhang, Jun-Zhong Zou, Lan-Lan Chen, Chunmei Wang, and Min Wang

Abstract In this paper, a new method to determine the penalty coefficients for different samples for the support vector machine (SVM) algorithm was proposed. Sequential minimal optimization (SMO) was then used to solve the SVM problem. Simulation results from applying the proposed method to binary classification problems show that the generalization error of the proposed method was smaller than standard SVM algorithm in the cases that the sizes of binary sample training sets (1) were selected in proportion; (2) were the same; (3) were quite different.

Keywords Support vector machine · Sequential minimal optimization · Structural risk minimization

1 Introduction

Support vector machine was proposed by Vapnik in the middle of 1990s [1]. It can be taken as a general machine learning algorithm based on limited samples theory. SVM is an effective method to avoid the local minimum and over-fitting problems that occur in conventional classification methods. Besides, it avoids the dimension disaster effectively by introducing a kernel function to analyze the nonlinear relationship in a higher dimensional space.

Denote the trainning sample set T as: $\{(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_m, y_m)\}$, where $\mathbf{x}_i \in \mathrm{R}^n$, $y_i \in \{1, -1\}, i = 1, ..., m$. To maximize the margin of the classifier, an optimization problem is formulated as follows:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \kappa\left(x_i, x_j\right) - \sum_{j=1}^{m} \alpha_j = \min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \tag{1}$$

J. Zhang (✉)
School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
e-mail: zhjmaster@sina.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_96,

$$s.t.\ \sum_{i=1}^{m} y_i \alpha_i = 0\,, \tag{2}$$

$$0 \le \alpha_{\mathrm{i}} \le \mathrm{C}, \mathrm{i} = 1, \cdots, \mathrm{m}. \tag{3}$$

Where $Q_{\mathrm{ij}} = y_i y_{\mathrm{j}} \kappa\left(x_{\mathrm{i}}, x_{\mathrm{j}}\right)$ is the kernel matrix; $e = (1, 1, \cdots, 1)^{\mathrm{T}}$; is the penalty coefficient. Gaussian kernel function $\kappa_{ij} = \exp\left(-\left\|x_i - x_j\right\| / 2\sigma^2\right)$ is used in this paper. After solving this optimization problem, we get the optimization resolution. One component from $\alpha^* = \left(\alpha_1^*, \cdots \alpha_{\mathrm{m}}^*\right)^{\mathrm{T}}$ is chosen and denoted as α^* . The threshold value is calculated as $\kappa_{ij} = \exp\left(-\left\|x_{\mathrm{i}} - x_{\mathrm{j}}\right\| / 2\sigma^2\right)$. and the decision function is computed as: $\mathrm{f}(x) = \mathrm{sgn}\left(\sum_{\mathrm{i}=1}^{\mathrm{m}} \alpha_{\mathrm{i}}^* \mathrm{y}_{\mathrm{i}} \kappa_{ij} + \mathrm{b}^*\right)$.

Conventional convex quadratic programming problems need to store the kernel matrix Q. This matrix is a two-dimensional m × m array. The size of matrix Q will grow quadratically with the increase of the sample size *m* and will saturate the memory space. Besides, a lot of training time is consumed in the matrix operations as m grows large. Osuna took the lead in proposing the decomposition strategy to solve the optimization problem and proved its convergence [2]. Platt introduced SMO Algorithm which was a particular case of decomposition strategy when the size of work sets is two [3]. Based on the SMO algorithm, Keerthi modified the Karush-Kuhn-Tucker (KKT) conditions and proposed two modified algorithms [4]. The convergence for both algorithms has been proved [5].

This paper is organized as follows: firstly, the existing problems of standard support vector machine are analyzed, and a new method is proposed to determine the penalty coefficient of different samples for the support vector machine algorithm. Secondly, an SMO algorithm is modified to solve this model. Finally, the performance of the proposed algorithm is evaluated by experiments.

2 Algorithms Design

2.1 Determination of C_i

In the standard SVM, each sample uses the same penalty coefficient *C*, and a larger value of *C* implies a small error rate. The determination of *C* is relevant to the location of samples. As shown in Fig. 1, there are two cases: (1) if a sample is surrounded by samples of the same class, the samples near this region should also be of same class, otherwise the error rate will increase, and this situation should be controlled by a lager value of *C*; (2) If a sample is surrounded by samples of a different class, the probability that the neighboring region is of the same class as the sample under evaluation is small, and the error rate will increase if the region expands, so this situation should be controlled by a smaller value of *C*. Each sample $\mathbf{x}_i$ has a separate Penalty coefficient. The model of standard support vector machine is modified as:

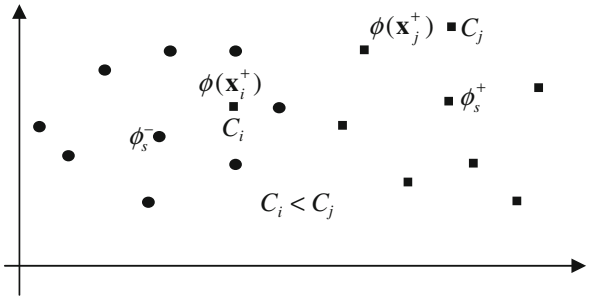


Fig. 1 Penalty coefficient of different samples

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \kappa \left(x_i, x_j\right) - \sum_{j=1}^{m} \alpha_j = \min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \tag{4}$$

$$s.t. \sum_{i=1}^{m} y_i \alpha_i = 0\,, \tag{5}$$

$$0 \le \alpha_{\mathrm{i}} \le \mathrm{C_i}, \mathrm{i} = 1, \cdots, \mathrm{m}. \tag{6}$$

In the feature space, $\mathbf{x}_i^+$ denotes the sample of class +1, $\mathbf{x}_i^-$ denotes the sample of class -1, ϕ_s^+ denotes the centroid of class +1, ϕ_s^- denotes the centroid of class -1, and l^+ is the number of positive samples, l^- is the number of negative samples, then:

$$\phi_s^+ = \frac{1}{l^+} \sum_{i=1}^{l^+} \phi(\mathbf{x}_i^+), \quad \phi_s^- = \frac{1}{l^-} \sum_{i=1}^{l^-} \phi(\mathbf{x}_i^-) \tag{7}$$

The distance from the sample of class +1 to the centroid of class -1 is:

$$\begin{aligned} d^+(\mathbf{x}^+) &= \left\| \phi(\mathbf{x}^+) - \phi_s^- \right\|^2 = \left\langle \phi(\mathbf{x}^+), \phi(\mathbf{x}^+) \right\rangle + \left\langle \phi_s^-, \phi_s^- \right\rangle - 2 \left\langle \phi(\mathbf{x}^+), \phi_s^- \right\rangle \\ &= \kappa(\mathbf{x}^+, \mathbf{x}^+) + \tfrac{1}{l^{-2}} \sum_{i,j=1}^{l^-} \kappa(\mathbf{x}_i^-, \mathbf{x}_j^-) - \tfrac{2}{l^-} \sum_{i=1}^{l^-} \kappa(\mathbf{x}^+, \mathbf{x}_i^-) \end{aligned} \tag{8}$$

Similarly, the distance from the sample of class -1 to the centroid of class +1 is:

$$d^-(\mathbf{x}^-) = \kappa(\mathbf{x}^-, \mathbf{x}^-) + \frac{1}{l^{+2}} \sum_{i,j=1}^{l^+} \kappa(\mathbf{x}_i^+, \mathbf{x}_j^+) - \frac{2}{l^+} \sum_{i=1}^{l^+} \kappa(\mathbf{x}^-, \mathbf{x}_i^+) \tag{9}$$

For positive samples, smaller value of d^+ implies that probability of the above case (2) is larger; on the contrary, lager value of d^+ implies that probability of the above case (1) is larger. So we can use following formula to calculate the penalty coefficient C_i:

$$C_i^+ = Cd^+(\mathbf{x}_i^+)/d_{\max}^+, \quad C_i^- = Cd^-(\mathbf{x}_i^-)/d_{\max}^- \tag{10}$$

Then $0 \le C_i^+, C_i^- \le C$.

2.2 Modification to the SMO Algorithm

To solve optimization problems of two lagrange multipliers, we need to identify the range of one multiplier first. Set C_1 and C_2 as the penalty coefficients of two multipliers α_1 and α_2, if $y_1 = y_2$, then $\alpha_1 + \alpha_2 = \gamma$ (γ is a constant). There are two situations as follows:

$$C_2 < C_1, \begin{cases} 0 < \gamma \le C_1, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = \gamma; \\ C_1 < \gamma \le C_2, & Min\ \alpha_2 = \gamma - C_1,\ Max\ \alpha_2 = \gamma; \\ C_2 < \gamma \le C_1 + C_2, & Min\ \alpha_2 = \gamma - C_1,\ Max\ \alpha_2 = C_2. \end{cases}$$

$$C_2 < C_1, \begin{cases} 0 < \gamma \le C_2, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = \gamma; \\ C_2 < \gamma \le C_1, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = C_2; \\ C_1 < \gamma \le C_1 + C_2, & Min\ \alpha_2 = \gamma - C_1,\ Max\ \alpha_2 = C_2. \end{cases}$$

If $y_1 \ne y_2$, then $\alpha_1 - \alpha_2 = \gamma$, there are two similar situations:

$$C_2 < C_1, \begin{cases} \gamma \le C_1 - C_2, & Min\ \alpha_2 = -\gamma,\ Max\ \alpha_2 = C_2; \\ C_1 - C_2 < \gamma \le 0, & Min\ \alpha_2 = -\gamma,\ Max\ \alpha_2 = C_1 - \gamma; \\ 0 < \gamma \le C_1, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = C_1 - \gamma. \end{cases}$$

$$C_2 < C_1, \begin{cases} -C_2 < \gamma \le 0, & Min\ \alpha_2 = -\gamma,\ Max\ \alpha_2 = C_2; \\ 0 < \gamma \le C_1 - C_2, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = C_2; \\ C_1 - C_2 < \gamma, & Min\ \alpha_2 = 0,\ Max\ \alpha_2 = C_1 - \gamma. \end{cases}$$

The remainder part of SMO algorithm is unchanged.

3 Experiment and Result Analysis

Take the data set pima-indians-diabetes[1] for example. This data set contains 768 samples and 8 characters altogether. In these 768 samples, there are 268 class +1 samples and 500 class −1 samples. Firstly, the sizes of binary sample training sets are randomly selected in three ways: (1) the sizes are in proportion; (2) the sizes are the same; (3) the sizes vary greatly. And in each way we select two sets of data. Secondly, the data set was standardized by $x_{ij}' = (x_{ij} - x_j)/s_j$, where x_{ij} is the jth character of the ith sample, S_j is the mean of the jth character, x_j is the standard deviation of the jth character. Gaussian kernel function was adopted. By searching, the maximum value of test correct rate is shown in the table below.

[1] http://mlearn.ics.uci.edu/databases/pima-indians-diabetes/

Table 1 Comparison between standard SVM and proposed SVM

Training sample number	Test sample number	Maximum value of test correct rate	
		Standard SVM (%)	Proposed SVM (%)
575(200+,375−)	193(68+,125−)	84.46	84.97
300(100+,200−)	468(168+,300−)	76.50	77.56
400(200+,200−)	368(68+,300−)	79.09	81.35
100(50+,50−)	668(218+,450−)	74.70	76.35
450(50+,400−)	318(218+,100−)	48.43	73.90
250(200+,50−)	518(68+,450−)	43.05	88.03

From Table 1, in the case that the sizes of training sample sets are selected in proportion, the generalization performance of standard SVM is closed to proposed SVM; in the case that the sizes of training sample sets are same, the generalization performance of proposed SVM is obviously superior to standard SVM; then in the case that the sizes of training sample sets are quite different, standard SVM loses the predictive capability, but proposed SVM still maintains a good generalization performance.

4 Conclusion

According to the principle of structural risk minimization, the generalization ability of learning machine is determined by VC dimension of a function set. By using different penalty coefficient $0 \leq C_i^+, C_i^- \leq C$, the diversity of function set is controlled effectively, we get a lower VC dimension learning machine, it has a better predictive capability. In this paper, the possibility of samples of different category appearing around a sample is considered. It is found that the possibility of the appearance of samples of the other category increases when the distance to the centriod of samples of the other category decreases. In future research work, the value of C_i will be calculated by the sample distribution around the sample $\mathbf{x}_i$.

References

1. Vapnik, V.N.: Statistical Learning Theory. New York, NY: Wiley (1995).
2. Osuna, E., Freund, R., Girosi, F.: Training support vector machines: An application to face detection. In: Plummer, D., ed.: Proceedings of the Compute Vision and Pattern Recognition. New York, NY: IEEE Press (1997) 130–136.
3. Platt, J.: Fast training of support vector machines using sequential minimal optimization. Advances in Kernel Methods—Support Vector Learning. Cambridge, MA: MIT Press (1999) 185–208.
4. Keerthi, S., Shevade, S., Bhattcharyya, C., et a1.: Improvements to Platt's SMO algorithm for SVM classifier design. Neural Comput. **13**(3) (2001) 637–649.
5. Keerthi, S., Gilbert, E.: Convergence of a generalized SMO algorithm for SVM classifier design. Mach. Learn. **46**(1/3) (2002) 351–360.

Recognition of Epileptic EEG Using Support Vector Machines

Chunmei Wang, Jun-Zhong Zou, Jian Zhang, Lan-Lan Chen, and Min Wang

Abstract A new approach based on support vector machine (SVM) is presented for the recognition of epileptic EEG. Firstly, the original signals of normal and epileptic EEG are decomposed with multi-resolution wavelet analysis. Secondly, their approximate entropy (ApEn) is estimated to extract features from raw EEG data. Finally, a SVM classifier with a Gaussian kernel function of SVM is used for the classification. Simulation results demonstrated that the SVM combined with wavelet transform and ApEn achieves high recognition accuracies.

Keywords Support vector machine · Epileptic EEG · Multi-resolution wavelet analysis · Approximate entropy

1 Introduction

The electroencephalogram (EEG) is one of the most common information sources used for studying brain function and neurological disorders. Different patterns of EEG are presented at the different physiological states. Epileptic seizure is the result of an occasional, sudden and excessive electrical discharge of the brain gray matter. The presence of epileptic activities (sharp wave, spike wave, spike-slow wave and sharp-slow wave) in the EEG supports the diagnosis of epilepsy. The International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) defined "spike" as a transient, with a duration from 20 to 70 ms (i.e. 1/50–1/15 s) and a "sharp wave" as a transient with similar characteristics but a duration of 70–200 ms (i.e. 1/14–1/5 s). Spike-slow wave and sharp-slow wave are

C. Wang (✉)
Department of Automation, East China University of Science and Technology, 200237 Shanghai, China; Department of Electronic Engineering, Shanghai Normal University, 200234 Shanghai, China
e-mail: wangchunmei01@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_97,

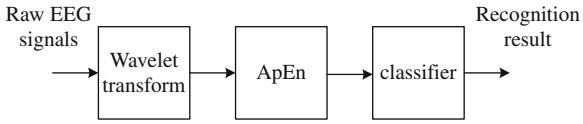


Fig. 1 Recognition of epileptic EEG based on SVM

complex discharges, which consist of a spike or sharp wave followed by a slow wave, respectively [1].

Recently, the support vector machines (SVM) learning method has become a very effective method for general purpose pattern recognition [2], which can handle the classification problems successfully. Classification is profoundly influenced by feature extraction. Since the EEG is non-stationary in general, it is more appropriate to use the time-frequency domain methods [3]. Then, ApEn values of different wavelet detailed coefficients are computed to obtain stable and credible values of ApEn. Figure 1 shows the structure of the recognition method based on SVM.

2 Support Vector Machine

The SVM proposed by Vapnik et al. [2] has been studied extensively for classification, regression and density estimation. Study demonstrates that SVM is an effective method not only avoid the local minimum and over-fitting problems that occur in conventional classification methods, but also solve the dimension curse by introducing kernel function, which makes it feasible to analyze the nonlinear relationship in a higher dimensional space.

Let $T = \{(x_1, y_1), \cdots, (x_{\mathrm{m}}, y_{\mathrm{m}})\} \in (\mathrm{X} \times \mathrm{Y})^{\mathrm{m}}$ be training samples set, where $x_i \in \mathrm{R}^{\mathrm{m}}$, $y_i \in \{1, -1\}$, $\mathrm{i} = 1, \cdots, \mathrm{m}$.

By introducing maximal interval method, an optimization problem is formed as follows:

$$\min_{\alpha} \frac{1}{2} \sum_{i,\mathrm{j}=1}^{m} y_i y_j \alpha_i \alpha_j \kappa(x_i, x_j) - \sum_{j=1}^{m} \alpha_j = \min_{\alpha} \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{Q} \boldsymbol{\alpha} - e^T \boldsymbol{\alpha} \tag{1}$$

where $C \geq \alpha_i \geq 0$, $\mathrm{i} = 1, \cdots, \mathrm{m}$; $\sum_{i=1}^{m} y_i \alpha_i = 0$, $Q_{i,\mathrm{j}} = y_i y_j \kappa(x_i, x_j)$ is the kernel matrix, C is a regularization parameter, $e = (1, 1, \cdots, 1)^{\mathrm{T}}$. Some typical kernel functions are polynomial kernel function and Gaussian kernel function.

3 Methods

3.1 Clinical Data

The EEG data required for this study consist of three different sets, which are collected at Changhai Hospital Attached to the Second Military Medical University, Shanghai, China. The first set includes surface recordings from four healthy

Table 1 Frequency distribution of three sub-bands at sample frequency of 128 Hz

Decomposed signal	Frequency band (Hz)
D1	32–64
D2	16–32
D3	8–16
A3	4–8

subjects. The data for the last two sets are obtained from five epileptic patients, which consists epileptic sharp wave and epileptic spike wave, respectively. Each data set consists of 100, 64 and 69 EEG epochs. The EEG data are digitized at 128 Hz sampling rate through 16-bit converter.

3.2 Processing of EEG Using DWT

Because the optimal resolution for analyzing epileptic activity corresponds to the frequency band 4–32 Hz [4], third level wavelet decomposition is utilized to EEG data for both normal subjects and epileptic patients. The corresponding frequency sub-bands of decomposed signals are listed in Table 1. Daubechies 4 (DB4) is chosen as a basis function.

3.3 Approximate Entropy (ApEn) Estimation

In this step, ApEn values of the detail coefficients (D1, D2 and D3) at each level of the wavelet decomposition are computed. ApEn has been used to characterize different biomedical signals [5, 6].

For computing the ApEn, we choose r=0.15*std and N=1,000, respectively based on suggestions by Pincus. Significant differences are found between the ApEn values of the first EEG data set (normal subjects) and the last two sets (patients with epilepsy). Moreover, another finding is that the ApEn values of different kinds of epileptic activities will have different distributions.

4 Experimental Results

We perform recognition based on SVM through two means.

Firstly, ApEn values for D1, D2 and D3 of each group are classified with one-dimensional input vector respectively. In group 1, 32 epileptic sharp samples and 50 normal EEG samples are used for training and the remaining is used for testing. In group 2, 35 epileptic spike samples and 50 normal EEG samples are used for training and the remaining is used for testing. Gaussian kernel function is chosen and classification results of two groups are presented in Table 2. The results indicate that optimal classification accuracy of epileptic sharp from normal EEG is in D2, and D1 of group 2 outperforms D2 and D3 in separating epileptic spike EEG from the normal EEG. The finding is consistent with the result we have got in our research work using Neyman-Pearson as a classifier.

Table 2 Classification accuracy of one-dimensional for two groups

Group	Sub-band	Parameter of SVM	Accuracy of training set (%)	Accuracy of test set (%)
1	D1	C=50, $2\sigma^2$=10	92.68	87.80
1	D2	C=50, $2\sigma^2$=10	97.56	95.12
1	D3	C=50, $2\sigma^2$=10	92.68	90.24
2	D1	C=20, $2\sigma^2$=10	96.43	94.05
2	D2	C=20, $2\sigma^2$=10	65.88	96.43
2	D3	C=20, $2\sigma^2$=10	76.47	89.29

Table 3 Classification accuracy of three-dimension for two groups

Group	Sub-band	Parameter of SVM	Classification accuracy (%)
1	D1,D2,D3	C=50, $2\sigma^2$=10	98.78
2	D1,D2,D3	C=20, $2\sigma^2$=10	98.81

Secondly, three-dimensional input vectors are formed by D1, D2 and D3 in each group. The 32*3 epileptic samples and 50*3 normal samples in groups are trained and the remaining is used for testing. Classification results of two groups are presented in Table 3.

Classification accuracies of three-dimensional input vectors are higher than the results of one-dimensional input vectors.

5 Conclusion and Future Work

Compared with clinical EEG data analysis, features related to the time-frequency and ApEn values distribution are extracted. Then a classifier based on SVM is applied. The results of EEG signal classification indicate that every different epileptic EEG has optimal classification accuracies in different sub-band. The SVM used for three-dimensional input vectors demonstrates SVM has strong generalization ability.

Further work can be performed to improve the classification accuracies by using different preprocessing and feature extraction methods.

References

1. IFSECN.: A glossary of terms commonly used by clinical electroencephalographers. Electroencephalogram. Clin. Neurophysiol. **37**(1974) 538–548.
2. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, Berlin, Heidelberg, New York, NY.
3. Khan, Y.U., Gotman, J.: Wavelet based automatic seizure detection in intracerebral electroencephalogram. Clin. Neurophysiol. **114** (2003) 898–908.
4. Subasi, A.: Automatic detection of epileptic seizure using dynamic fuzzy neural networks. Expert Syst. Appl. **31** (2006) 320–328.

5. Pincus SM (2001) Assessing serial irregularity and its implications for health. Ann N Y Acad Sci **954** (2001) 245–267.
6. Ocak, H.: Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Systems with Applications (2008), doi:10.1016/j.eswa.2007.12.065

Image Denoising Using Noisy Chaotic Neural Networks

Leipo Yan and Lipo Wang

Abstract This paper uses the noisy chaotic neural network (NCNN) that we proposed earlier for image denoising as a constrained optimization problem. The experimental results show that the NCNN is able to offer good quality solutions.

Keywords Image processing · Neural networks · Chaos

1 Introduction

The objective of image denoising is to estimate the original image from a noisy image with some assumptions or knowledge of the image degradation process. There exist many approaches for image denoising [1, 2]. Here we adopt a Bayesian framework because it is highly parallel and it can decompound a complex computation into a network of simple local computations [2], which is important in hardware implementation of neural networks. This approach computes the maximum a posteriori (MAP) estimation of the original image given a noisy image. The MAP estimation requires the prior distribution of the original image and the conditional distribution of the data. The prior distribution of the original images imposes contextual constraints and can be modeled by Markov random field (MRF) or, equivalently, by Gibbs distribution. The MAP-MRF principle centers on applying MAP estimation on MRF modeling of images.

Li proposed the augmented Lagrange Hopfield method to solve the optimization problem [3]. He transformed a combinatorial optimization problem into real constrained optimization using the notion of continuous relaxation labeling. The HNN was then used to solve the real constrained optimization.

The neural network approaches have been shown to be a powerful tool for solving the optimization problems [4]. The HNN is a typical model of neural network

L. Yan (✉)
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore
e-mail: elpwang@ntu.edu.sg

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_98,

with symmetric connection weights. It is capable of solving quadratic optimization problems. However, it suffers from convergence to local minima. To overcome the weakness, different simulated annealing techniques have been combined with the HNN to solve optimization problems [5]. The TCNN showed good performance in solving traveling salesman problem [6] . However CSA is deterministic and is not guaranteed to settle down at a global minimum. In view of this, Wang and Tian [7] proposed a novel algorithm called stochastic chaotic simulated annealing (SCSA) which combines both stochastic manner of SSA and chaotic manner of CSA. In this paper the NCNN, which performs SCSA algorithm, is applied to solve the constrained optimization in the MAP-MRF formulated image denoising. Experimental results show that the NCNN outperforms the HNN and the TCNN.

2 The Noisy Chaotic Neural Network for Image Denoising

Let $u_i(I)$ denote the internal state of the neuron (i, I) and $p_i(I)$ denote the output of the neuron (i, I). $p_i(I) \in [0, 1]$ represents the strength that the pixel at location i takes the value I. The NCNN is formulated as follows [7]:

$$p_i^{(t)}(I) = \frac{1}{1 + e^{-u_i^{(t)}(I)/\epsilon}} \tag{1}$$

$$\begin{aligned} u_i^{(t+1)}(I) = {} & k u_i^{(t)}(I) - z^{(t)}(p_i^{(t)}(I) - I_o) + n^{(t)} + \\ & \alpha\left(\sum_{i'=1, i' \neq i}^{N} \sum_{I'=1}^{M} T_{iI;i'I'} p_{i'}^{(t)}(I') + \mathcal{I}_i(I)\right) \end{aligned} \tag{2}$$

$$z^{(t+1)} = (1 - \beta_z) z^{(t)} \tag{3}$$

$$A[n^{(t+1)}] = (1 - \beta_n) A[n^{(t)}] \tag{4}$$

where

$T_{iI;i'I'}$: connection weight from neuron (i', I') to neuron (i, I);
$\mathcal{I}_i(I)$: input bias of neuron (i, I);
k: damping factor of nerve membrane $(0 \leq k \leq 1)$
α: positive scaling parameter for inputs;
ϵ: steepness parameter of the output function $(\epsilon \geq 0)$;
z: self-feedback connection weight or refractory strength $(z \geq 0)$;
I_o: positive parameter;
n: random noise injected into the neurons;
β_z : positive parameter $(0 < \beta_z < 1)$;
β_n : positive parameter $(0 < \beta_n < 1)$;
$A[n]$: the noise amplitude.

When $n^{(t)}$ in (2) equals to zero, the NCNN becomes TCNN. When $z^{(t)}$ equals to zero, the TCNN becomes similar to the HNN with stable fixed point

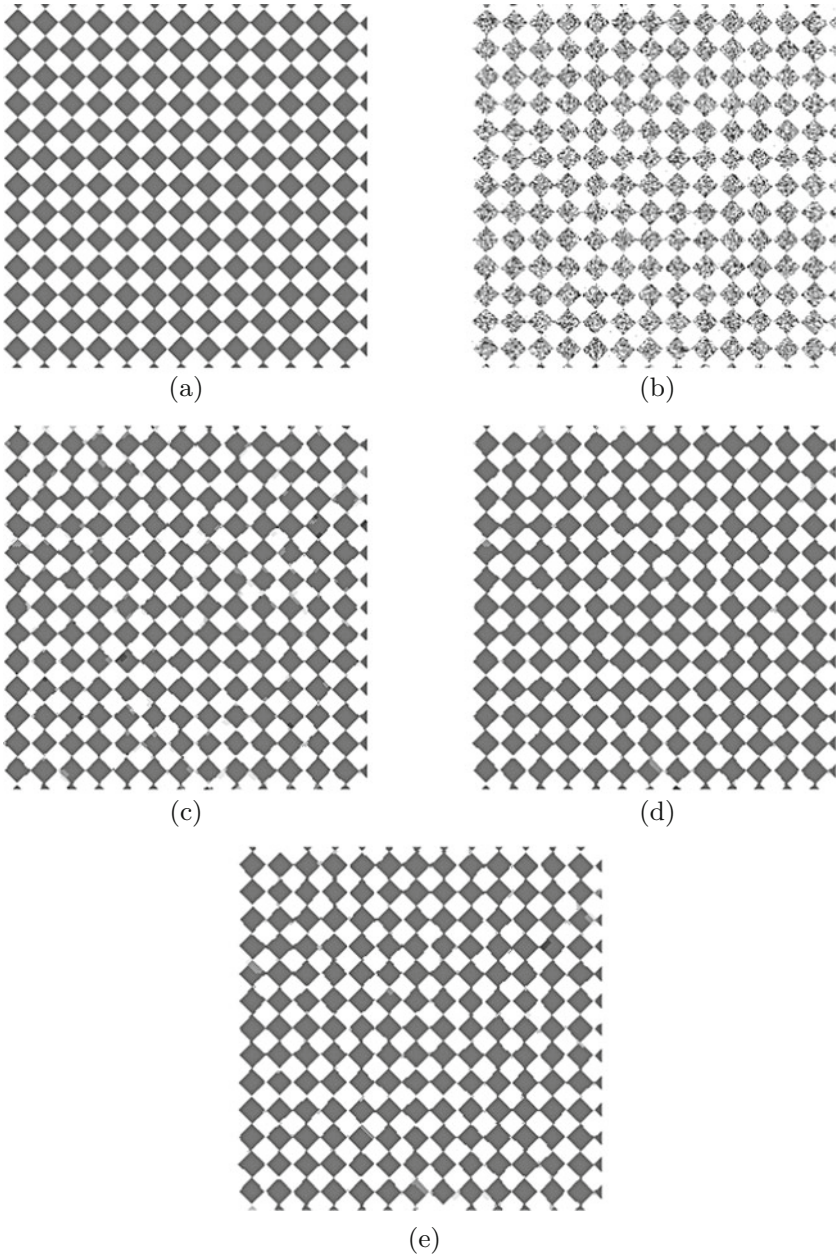


Fig. 1 Denoising of the square image with noise level $\sigma = 0.5$: (**a**) Original image. (**b**) Noisy image. (**c**)–(**e**) Denoised images using the HNN, the TCNN and the NCNN, respectively

dynamics. The basic difference between the HNN and the TCNN is that a nonlinear term $z^{(t)}(p_i^{(t)}(I) - I_o)$ is added to the HNN. Since the "temperature" $z^{(t)}$ tends toward zero with time evolution, the updating equations of the TCNN eventually reduce to those of the HNN. In (2) the variable $z^{(t)}$ can be interpreted as the strength of negative self-feedback connection of each neuron, the damping of $z^{(t)}$ produces successive bifurcations so that the neurodynamics eventually converge from strange attractors to a stable equilibrium point [5].

CSA is deterministic and is not guaranteed to settle down to a global minimum. In view of this, Wang and Tian [7] added a noise term $n^{(t)}$ in (2). The noise term continues to search for the optimal solution after the chaos of the TCNN disappears.

The dynamics of the NCNN for image denoising is described by (more details will be discussed elsewhere):

$$u_i^{(t+1)}(I) = ku_i^{(t)}(I) - z^{(t)}(p_i^{(t)}(I) - I_o) + n^{(t)} + \alpha q_i^{(t)}(I) \tag{5}$$

where

$$q_i(I) = -\frac{\partial L_\beta(p, \gamma)}{\partial p_i(I)}$$

Note that the Lagrange multipliers are updated with neural outputs according to $\gamma_k^{(t+1)} = \gamma_k^{(t)} + \beta C_i(p^{(t)})$. Experimental results are shown in Fig. 1.

3 Conclusion

We have implemented the noisy chaotic neural network (NCNN) to address the MAP-MRF formulated image denoising problem. NCNN effectively overcomes the local minimum problem because of its SCSA characteristics. We have shown that the NCNN gives better quality solutions compared to the HNN and the TCNN.

References

1. Andrews, H.C., Hunt, B.R.: Digital Image Restoration. Englewood Cliffs, NJ: Prentice-Hall (1977).
2. Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions and the bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. **6**(6) (1984) 721–741.
3. Wang, H., Li, S.Z., Petrou, M.: Relaxation labeling of Markov random fields. IEEE Trans. Syst. Man Cybern. **20** (1990) 709–715..
4. Hopfield, J.J., Tank, D.W.: Neural computation of decisions in optimization problems. Biol. Cybern. **52** (1985) 141–152.
5. Chen, L., Aihara, K.: Chaotic simulated annealing by a neural network model with transient chaos. Neural Netw, **8** (6) (1995) 915–930.
6. Nozawa, H.: A neural network model as a globally coupled map and applications based on chaos. Chaos **2**(3) (1992) 377–386.
7. Wang, L., Tian, F.: Noisy chaotic neural networks for solving combinatorial optimization problems. Proceedings of the International Joint Conference on Neural Networks, vol. 4 (July 2000) 4037–4040.

Increase Productivity in Circuit Board Manufacturing Through Support Vectors

Hai Tuyen Tran and Chia-Jiu Wang

Abstract This paper investigates the possibility of using support vector machines (SVM) to increase productivity in manufacturing printed circuit boards. The current method of setting up a machine to place solder pastes on circuit boards is a very sophisticated and time consuming process, requiring many trial and error steps. SVMs are used to predict the outcome of a new production run based on support vectors extracted from machine setup parameters used in previous production runs. By incorporating a SVM machine in the production line, the amount of solder to be placed on a circuit board can be effectively monitored and adjusted at the early stage of the production line, resulting in much higher productivity of a SMT-based production line.

Keywords Support vectors · Support vector machines · Surface mount technology · SMT processes

1 Introduction

Manufacturing circuit boards require knowledge, investment and substantial experience. Due to the demand of shorter time to market, processes that produce circuit board need to be robust and outcomes need to be predictable. A Support Vector Machine (SVM) can be used to extract support vectors from data sets of an existing production line using Surface Mount Technology (SMT) which is trial and error adjusted by experienced engineers. With a SVM incorporated in a SMT-based production line, we can predict the results for a particular machine setting before starting operation of the production line. Support vector machine learning has been studied and reported by many researchers in many unique applications

H.T. Tran (✉)
Department of Electrical and Computer Engineering, University of Colorado, Colorado Springs, CO 80918, USA

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_99,

such as character recognition, text classification, underwater object detection and battery state of charge estimation [1–5]. In Section 2 a short description of surface mount technology based production line is presented. In Section 3, a modified SMT production line incorporating a SVM machine is discussed and the features of the production line are selected as inputs for SVM machines. Experimental results and discussions are presented in Section 4. The effectiveness of using SVMs in solder placing in a SMT-based production line has been demonstrated in this paper.

2 Surface Mount Technology Production Line Components

Due to the complexity of the circuit board used today, production lines using surface mount technology (SMT) is preferable over those using Pin Through Hole (PTH) technology. SMT processes can not only produce circuit boards at a faster rate, but also require less human interaction. A typical SMT-based production line contains a set of machines as depicted in Fig. 1

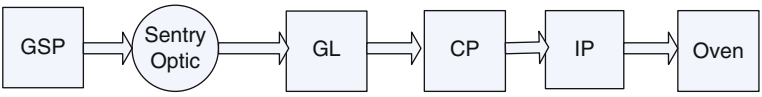


Fig. 1 A SMT-based production line containing six machines

A typical SMT based production line has six machines; GSP, CentryOptic, GL, CP, IP and Oven. GSP is a solder placing machine, applying solder through the stencil opening onto the circuit board. Each GSP has two blades to apply solder to the circuit boards. The speed of the blade, pressure of the blade and the table height of the GSP machine affect the amount of solder to be printed on the board. The amount of solder can be adjusted by the following settings: the height of the table, the speed of the blades or the pressure of the blades. Sentry Optic is a high resolution camera used in measuring the height of the solder paste. The data measured from the circuit board under test are compared with the pre-set data to determine whether the board passes the test or not. GL is the glue machine that places glue on heavy parts to prevent them from falling off in the next process. CP is used to place small devices on to the circuit board. Resistors, capacitors, inductors, and small ICs are placed on the circuit board by the CP. The CP has 20 heads and each head contains 6 nozzles. Depending on the side of the device, a different nozzle is chosen. The IP machine is also used in placing parts on the circuit board. The major difference between IP and CP is that IP places larger ICs and BGA-type parts. The last machine used in the production line is the oven which provides heat to reflow the solder. Each production line will have a separate heat profile. Then the circuit board is visually inspected by an operator. Considering all processing steps involved in make the circuit board, the outcome at the GPS mostly determines the final quality of the board. Since the output, i.e., the solder, of the GPS is measured by Sentry Optic, the board can be

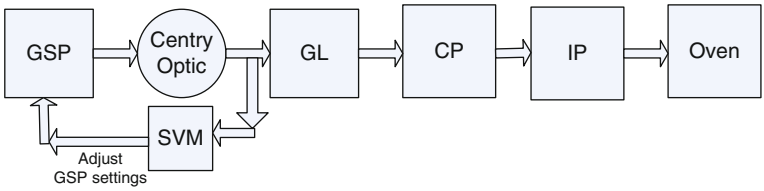


Fig. 2 The SMT-based production line with added a SVM machine

stopped from moving to the next machine if some discrepancies are detected in the solder.

With the current manufacturing process, engineers have to rely on past experience to set up the GSP and generally a trial and error method is adopted for a new production line. Here we propose a modified SMT-based production line as shown in Fig. 2 to integrate many engineers' experience into a SVM machine.

3 SVM Used in SMT-Based Production Line

The SVM machine can be used for classification and nonlinear prediction. Using the support vector learning algorithm, support vectors can be extracted from the data set collected at the Sentry Optic site. Then the SVM machine can use the support vectors and the current GSP setting to predict the outcome of the solder in terms of height which is the output of the GSP machine. The SVM will output a pass/fail signal to indicate that the board will pass or fail at the end of the production line. This SVM is used for failure detection on the height of the solder. If the board has a flag of failure, then it should be removed from the production line. The configuration of the SVM is shown in Fig. 3.

Hundreds of GSP settings and the corresponding solder height measured by the Sentry Optic are recorded. A pre-processing program extracts the speed of the blade, the pressure of the blade and the stencil table height. Another different pre-processing program is written to extract the height of the solder for a location specified by user input. Each data entry has three input components from GSP, i.e., speed, pressure, and height and one component from Optic Sentry, i.e., pass (1) or fail (-1). One example of the data entry is $\{2, 6, 2.6, 1\}$ for a pass, $\{2, 6, 2.5, -1\}$ for a fail.

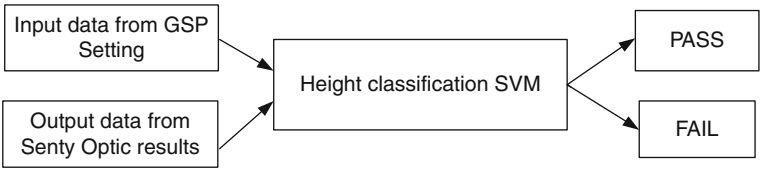


Fig. 3 The SVM configuration

MySVM written by Stefan Ruping is used in this work [6]. The following table shows the number of support vectors (SV) versus the capacity C of the SVM. Table 1

In our work we study C=1 as well as C=20. The input function selected is a polynomial of power 4 which is $K(X_i, X_j) = (1 - X_i^T X_j)^4$. Some testing results are shown in Figs. 4 and 5. It demonstrates that SVM predicts correctly.

Table 1 SVM Capacity versus number of support vectors

C	1	5	20	50	100	200	500
SV	16	16	15	15	15	15	15

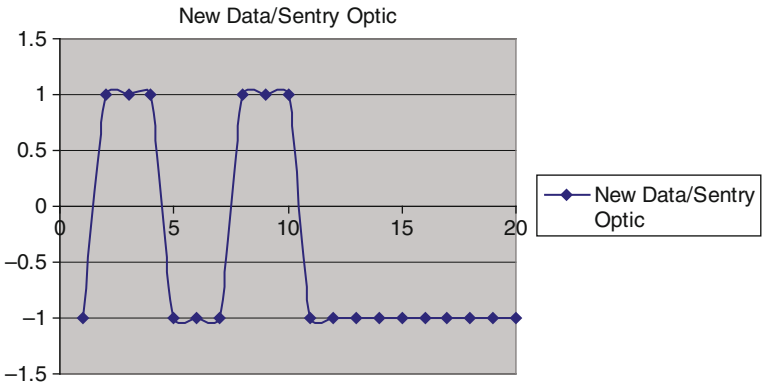


Fig. 4 Testing results using new GSP setting data

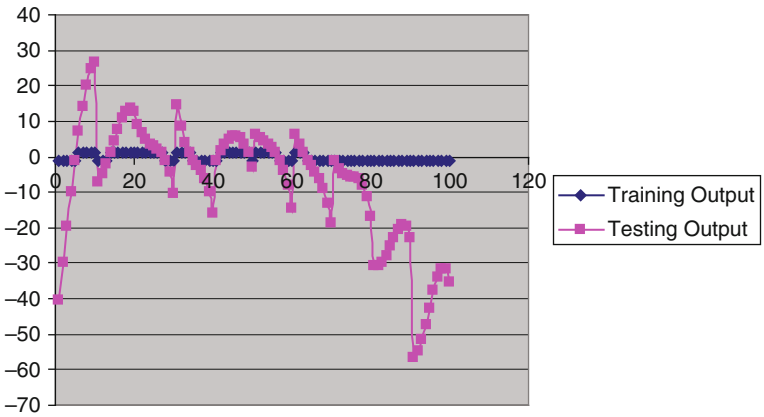


Fig. 5 Some SVM testing results

4 Discussions

Results shown in Section 3 show that a SVM machine can be used to predict a pass or fail for a solder placing machine such as a GSP machine. By using SVM, when using a new GSP setting for a new production line one can predict the solder height before the actual placement. Using the SVM, it helps engineers to reduce the machine setup time significantly during preparation of the production line. After the production is in full operation, a pass or fail signal generated by the SVM machine at the second stage can alarm engineers any variations occurred during the operation.

The same SVM method used in the SMT-based production line can be applied to other manufacturing tests or processes. For example, some particular measurements at Integrated Circuit Test (ICT) can be used to predict output at the functional test. Or a shorter ICT can be used to predict whether to run a long ICT test. These savings in time can add up and becomes significant throughout the manufacture process.

Acknowledgement Special thanks to Stefan Ruping, who created the mysvm software. Without the software this project will not exist. Special thanks to the management team at Sanmina-SCI for their help getting data for training and predicting for this project.

References

1. Vapnik, V.N.: The Nature of Statistical Learning Theory. New York, NY: Springer (1995).
2. Hansen, T., Wang, C.-J.: Support vector based battery state of charge estimator. J. Power Sources **141** (2005) 351–358.
3. Mitra, V., Wang, C.-J.: Content based audio classification: A neural network approach. Soft Comput **12**(7) (2008) 639–646.
4. Mitra, V., Wang, C.-J.: Satarupa Banerjee, "Lidar detection of underwater objects using a neuro-SVM based architecture". IEEE Trans. Neural Netw. **17**(4) (May 2006) 717–731.
5. Mitra, V., Wang, C.-J.: Content based audio classification: A neural network approach. Soft Comput. **12**(7) (2008) 639–646.
6. Ruping, S.: mySVM Manual.

Part XI
Emerging Technologies for Brain Computer Interfaces

Towards Affective BCI/BMI Paradigms – Analysis of fEEG and fNIRS Brain Responses to Emotional Speech and Facial Videos

Tomasz M. Rutkowski, Qibin Zhao, Andrzej Cichocki, Toshihisa Tanaka, and Danilo P. Mandic

Abstract A previous study in the visual domain showed that similar spatial patterns of electroencephalography (EEG) power can be seen when subjects are visually presented with short videos of faces expressing emotions. As a corollary study we examine both the auditory and visual expressions of the same emotions to understand how listeners process auditory-only and visual-only expressions of affective stimuli in functional EEG (fEEG) and functional near-infrared-spectroscopy (fNIRS) signals. For this purpose we utilize fEEG directed transfer functions (DTF) analysis in very short windows of 250 ms which potentially is of interest of brain-computer/machine-interfacring (BCI/BMI) application with utilization of affective (limbic) responses from human brains. The preliminary results with affective paradigms confirm the hypothesis of cortical emotional responses in fEEG and fNIRS.

Keywords Brain computer interface · Brain emotional responses · Neurophysiological signal processing · Functional EEG and NIRS

1 Introduction

Recent advances in brain-computer/machine-interfacing (BCI/BMI) have shown an urgent necessity to search for new and more user-friendly paradigms which would allow more natural interaction of humans and machines with utilization of so revealed new communication channels [1]. BCI/BMI related literature recognizes two general classes of BMI paradigms [2], those which are related to external environment stimuli, thus utilizing stimuli driven brain responses, and the other ones which are completely independent from environmental stimulation, thus relaying only on internal (imagery) brain activity managed by the users will. In this paper we focus on the first class of dependent and stimuli driven paradigms with

T.M. Rutkowski (✉)
Brain Science Institute RIKEN, Wako-shi, Saitama, Japan
e-mail: tomek@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_100,

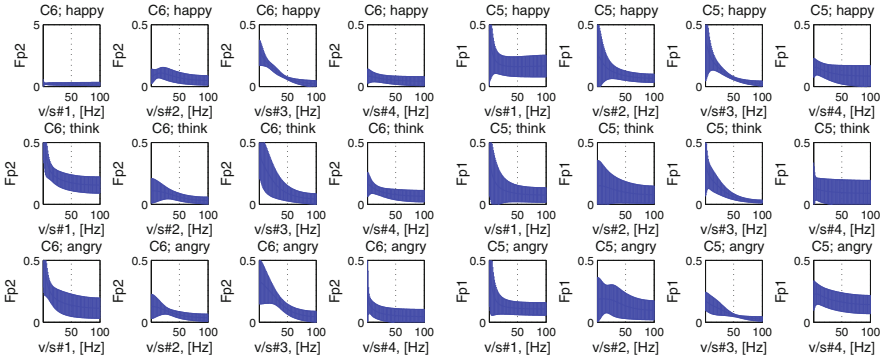
a concept of limbic (affective/emotional) responses involved in order to search for a new class of BCI/BMI paradigms which would involve very fast human–interface interaction based on affective responses. A curious question has been what is happening in the brain when people are perceiving or processing (e.g. imaging) expressions of emotions. It is known that the amygdala is a control center for emotions in humans [3], however, recent research has suggested that the frontal cortical areas are also involved in perception of emotions [4, 5]. Also, a recent work suggests that there is no single cortical brain area specifically involved in emotional processing, but there are neural hubs that process information from different parts of the brain [3]. Previous research related to emotional responses analysis, confirmed also by the authors [4], shown an asymmetry in the frontal cortex, depending on the valence of the emotion perceived. In that study stimuli was presented to listeners in form of video presentations (short 2–5 s long movies) of emotional facial expressions taken from the multimodal emotional corpus [6]. The emotions were happy, thinking and angry. Four subjects were asked to view the short movies on the computer screen, try not to blink, and then answer the question that appeared on the screen afterward, asking what emotion they thought the actor/actress was expressing. The results indicated that frontal cortical activity varies according to the particular emotion – happy and thinking showed asymmetry in cortical activity, with left cortical activity stronger for happy and right, stronger for thinking.

A question we want to address in this study is what happens when similar emotional audio and video stimuli are presented to subjects separately with fEEG electrodes over their temporal and frontal lobes, together with frontal head fNIRS activity monitoring. In the previous study also longer window power patterns were analyzed not allowing a simple BCI/BMI application. In the current study only very short 250 ms windows are analyzed for their fEEG spectral information flow between electrodes.

2 Methods

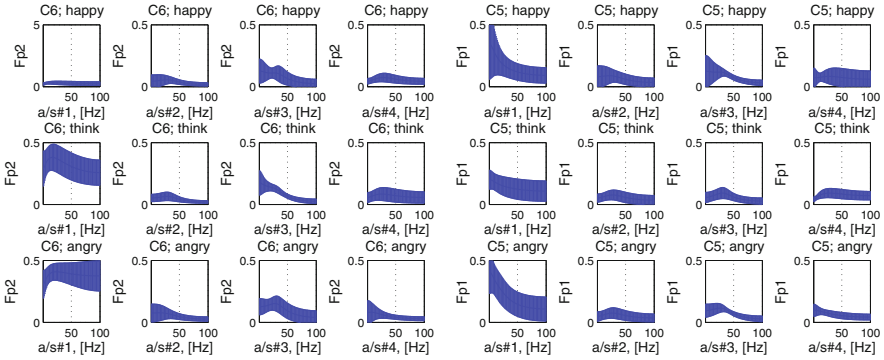
The fEEG experiment has been performed at the Advanced Brain Signal Processing Laboratory, RIKEN Brain Science Institute, using a g.USBamp fEEG data acquisition system with four electrodes placed over temporal lobes in positions $C5$, $C6$, $Fp1$ and $Fp2$. Additionally the fNRIS sensor over front head was placed to analyze cortex's blood flow activity in response to the stimuli. fEEG was recorded with sampling frequency of 256 Hz with impedance below 5 kΩ. The fNIRS was sampled with 6 Hz sampling frequency and only the dynamics of oxygenated hemoglobin signal (dO2Hb) was analyzed.

Data from four subjects as in Fig. 1 and 2 was analyzed. The subjects observed and listened five British English professional actors/actresses from the emotional multimodal corpus [6]. The affective speech and facial expressions examined were short utterances expressing happy, anger, and thinking. We compare the results



(a) EEG DTF($F_p2, C6$) - facial videos.

(b) EEG DTF($Fp1, C5$) - facial videos.



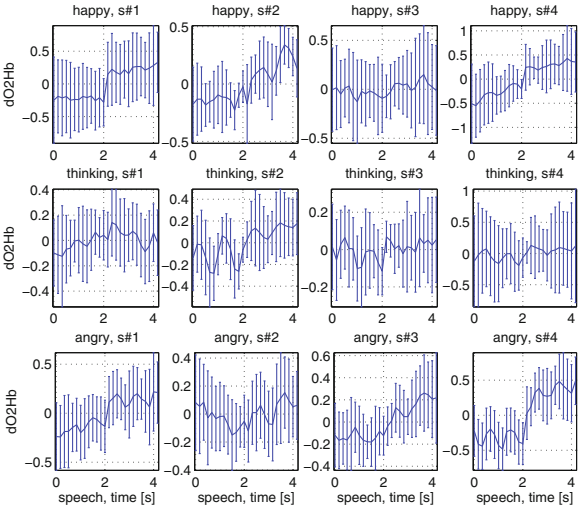
(c) EEG DTF($Fp2, C6$) - speech.

(d) EEG DTF($Fp1, C5$) - speech.

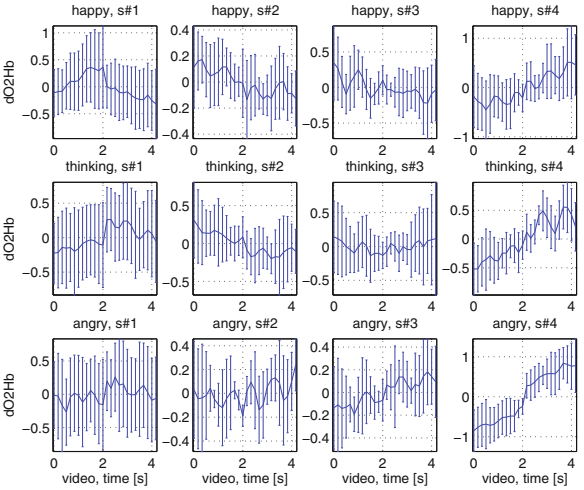
Fig. 1 DTF patterns for both affective stimuli conditions and two electrodes pairs $Fp2$, $C6$ and $Fp1$, $C5$

of the auditory and visual domain with those of the previous results in the visual domain only [4].

The fEEG signals are first preprocessed with a multichannel EMD extension as discussed in [7]. Next, directed transfer function (DTF) is calculated, as first proposed in [8], for the pair of electrodes $Fp1/C5$ and $Fp2/C6$, as $\gamma_{ij}^2(f) \triangleq \frac{|H_{ij}(f)|^2}{\sum_{j=1}^{m}|H_{ij}(f)|^2} \in [0,1]$, where the frequency-dependent normalization is performed so that $\gamma_{ij}^2(f)$ quanties the fraction of inflow to channel i stemming from channel j. DTF allows for evaluation of information transfer in different frequency sub-bands based on multivariate autoregressive model analysis of fEEG channels as presented in Fig. 1b, c, where mean values for six experiments for each subjects are summarized in form of mean values plots together with error bars.



(a) Emotional speech.



(b) Emotional facial videos.

Fig. 2 Frontal fNIRS for speech/audio (**a**) and faces/video (**b**) respectively. The transition point between *no stimuli* and *stimuli* was at 2 s, where significant changes in fNIRS mean values as well much smaller error bars are present

3 Conclusions

As a result of the presented analysis, we confirm our hypothesis of a possibility to quantify affective/emotional responses from fEEG as shown mostly for DTF analysis pairs *Fp*2, *C*6 for video and speech utterances. The mean DTF spectral responses together with standard deviation error bars are presented in Fig. 1a, c. The patterns for fEEG electrodes pairs *Fp*1, *C*5 did not show significant difference among the stimuli conditions as can be seen in Fig. 1b, d. The fNIRS results analysis shown

only interesting patterns at the pre-stimuli/stimuli transition points, were sudden and coherent (very small error bars) changes are visualized in Fig. 2a, b. The presented approach is a step forward in a search for new affective BCI approaches.

Acknowledgments The research presented in this paper was supported in part by the Japan Society for the Promotion of Science grant no. 21360179.

References

1. Cichocki, A., Washizawa, Y., Rutkowski, T., Bakardjian, H., Phan, A.H., Choi, S., Lee, H., Zhao, Q., Zhang, L., Li, Y.: Noninvasive BCIs: Multiway signal-processing array decompositions. Computer **41**(10) (2008) 34–42.
2. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain–computer interfaces for communication and control. Clin. Neurophysiol. **113** (2002) 767–791.
3. Pessoa, L.: On the relationship between emotion and cognition. Nat. Rev. Neurosci. **9** (2008) 148–158.
4. Rutkowski, T.M., Cichocki, A., Ralescu, A.L., Mandic, D.P.: Emotional states estimation from multichannel EEG maps. In Wang, R., Gu, F., Shen, E., eds.: Advances in Cognitive Neurodynamics ICCN 2007 Proceedings of the International Conference on Cognitive Neurodynamics. Neuroscience, Springer, Berlin, Heidelberg (2008) 695–698.
5. Lehrer, J.: How We Decide. Boston, MA, New York, NY: Houghton Mifflin Harcourt Publishing Company, (2009).
6. Baron-Cohen, S.: Mind Reading – The Interactive Guide to Emotions. London: Jessica Kingsley Publishers, (2004).
7. Rutkowski, T.M., Cichocki, A., Tanaka, T., Mandic, D.P., Cao, J., Ralescu, A.L.: Multichannel spectral pattern separation – an EEG processing application. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2009), Taipei, Taiwan, IEEE Press (2009) 373–376.
8. Kaminski, M., Blinowska, K.: A new method of the description of the information flow in the brain structures. Biol. Cybern. **65**(3) (1991) 203–210.

Imagery Movement Paradigm User Adaptation Improvement with Quasi-movements Phenomenon

Hiroshi Higashi, Tomasz M. Rutkowski, Yoshikazu Washizawa, Toshihisa Tanaka, and Andrzej Cichocki

Abstract We discuss a novel movement imagery brain-computer/ma-chine-interface (BCI/BMI) paradigm learning procedure with utilization of real- and quasi-movements of subjects' thumbs. In the proposed procedure volitional movements are slowly minimized by the subjects to a very low level so that finally they become undetectable by objective measures such as electromyography (EMG). The procedure allows the subjects to understand motion imagery process, which follows after the training. The procedure allows also to control the final movement imagery protocol and to detect any possible movements in case subject would not learn to suppress them completely. We present also a discussion on electroencephalography (EEG) signals pre-processing steps with common spatial pattern (CSP) method improvements. Promising results were obtained with subjects who could not perform the motion imagery paradigm as well with those who never tried it before conclude the paper.

Keywords Brain computer interface · Movement imagery paradigm · BCI user adaptation · Quasi movement paradigm

1 Introduction

Movement imagery belongs to a class of independent BCI/BMI paradigms relaying only on subjects' conscious control of brain dynamics captured later on scalp level in form of EEG [1–3]. A process of user learning to perform movement imagery tasks can sometimes take a long time [4] and this is why a "quasi-movement" paradigm was first proposed in [5]. Quasi-movement of a muscle is performed in a condition, when a voluntary movement cannot be detected by objective measures as EMG electrodes attached to skin over such muscle. Subjects usually train to

H. Higashi (✉)
Tokyo University of Agriculture and Technology, Tokyo, Japan
e-mail: higashi@sip.tuat.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_101,

perform such movement in a bio-feedback setup where they can watch amount of EMG power and minimize it by suppressing the muscle activity. Building on those reports we propose a substantial improvement of movement imagery paradigm with utilization of stepwise learning procedure starting from real movements, followed with quasi-movements with EMG biofeedback and finally with an imagery session only. We also propose a substantial reduction of EEG electrodes number to 12 only, comparing to previous reports [4, 5]. In real-, quasi- and imagery-movement paradigms amplitude dynamics of electroencephalographic *alpha* oscillations relate the regional involvements of cortical areas during the experimental task. In all the above conditions, a significant variations within *alpha* (so called *mu*-rythm) oscillations over the sensorimotor and motor areas of the contralateral hemisphere are usually detected [6].

2 Methods

The experiments are conducted in Advanced Brain Signal Processing Laboratory, RIKEN Brain Science Institute, Saitama, Japan. The five subjects during the experimental procedures are wearing EEG caps connected to g.USBamp bio-amplifiers capturing 12 EEG signals from electrodes located at $C1$, $C2$, $C3$, $C4$, $C5$, $C6$, $FC1$, $FC2$, $FC3$, $FC4$, $FC5$, $FC6$ as in international extended 10/20 system [6], with a sampling frequency of 256 Hz and impedance not exceeding 10 kΩ. The two pairs of bipolar EMG biofeedback electrodes for real- and quasi-movement muscle control are placed on the both subject hands to monitor activity of *an adductor pollicis* muscle.

A classification of left/right real-, quasi and imagery-movements based on EEG signals is performed in the presented approach. Recorded signals consist of $N_e = 12$ EEG channels and there are two target classes: *Class1* (left hand) and *Class2* (right hand). For each movement only $M = 10$ trials are performed. The EEG signals are filtered to fit a range 8 Hz to f_uHz frequency bands with a Butterworth filter order of 4. Let $\mathcal{X} = \{X_1, X_2, \ldots, X_{2M}\}$ where $X_i \in \mathbb{R}^{N_e \times N}$ denotes the filtered ith trial of EEG signal and N is the number of samples in the signal (640 samples stand for 2.5 s in this study). We adopt in the study a common spatial patterns (CSP) classification method [7]. The main idea of CSP is to determine spatial filters which maximize difference in variance. The cost function of CSP is expressed as

$$J[w] = \frac{w^T \mathbf{\Sigma}_1 w}{w^T (\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2) w}, \tag{1}$$

where $w \in \mathbb{R}^{N_e}$ is the unknown weight vector of the spatial filter. $\mathbf{\Sigma}_1$ is given by

$$\mathbf{\Sigma}_1 = \frac{1}{M} \sum_{i \in C_1} \frac{\boldsymbol{X}_i \boldsymbol{X}_i^T}{trace(\boldsymbol{X}_i \boldsymbol{X}_i^T)}, \tag{2}$$

where C_1 represents all *Class1* of EEG trials, *trace*$(\cdot)$ takes the sum of the diagonal elements, and $\boldsymbol{\Sigma}_2$ is also obtained in the same procedure. This cost function can be solved by the eigenvalue decomposition. Then we can obtain the two spatial filters, $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ by minimizing two cost funtions which have the numerator of $w_{\mathbf{1}}{}^T \boldsymbol{\Sigma}_{\mathbf{1}} w_{\mathbf{1}}$ and $w_{\mathbf{2}}{}^T \boldsymbol{\Sigma}_{\mathbf{2}} w_{\mathbf{2}}$ respectively. When a test signal, $\boldsymbol{X}$ is given, we classify the test signal by the following procedure

$$\begin{cases} \text{var}(w_1^T \boldsymbol{X}) < \text{var}(w_2^T X) \Rightarrow \boldsymbol{X} \in C_1, \\ \text{var}(w_2^T X) < \text{var}(w_1^T X) \Rightarrow X \in C_2, \end{cases} \tag{3}$$

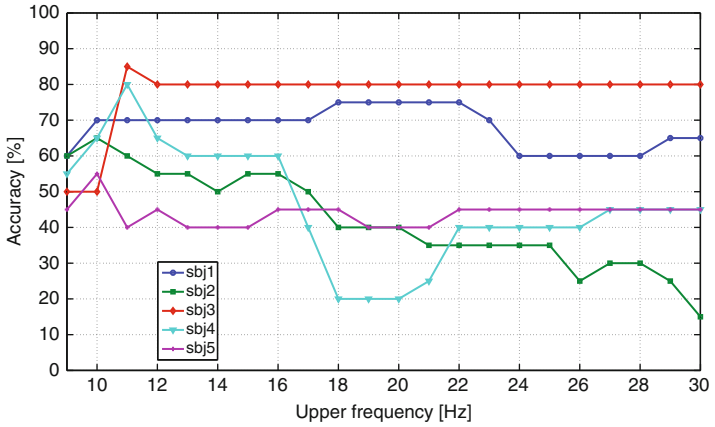
where var$(\cdot)$ is a variance.

3 Results and Conclusions

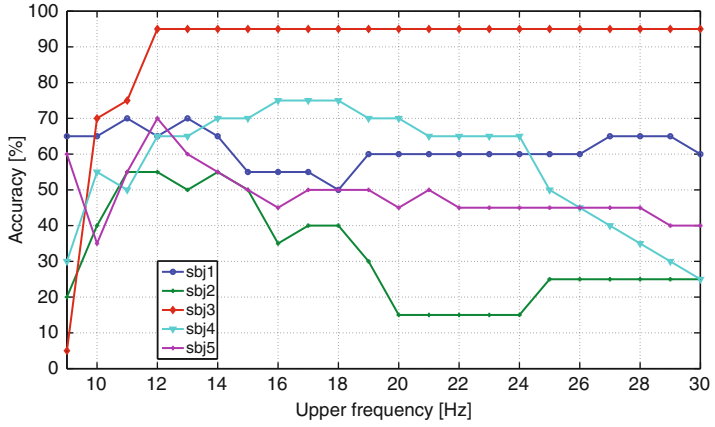
Figure 1a–c present the classification accuracy for each frequency band f_u. The classification accuracy is obtained by a leave-one-out cross validation method. Figure 1a presents results of EEG classification when users are actually moving their thumbs. This step is very useful for subjects' later imagery-movement paradigm execution. The classification accuracy is already high for most of the subjects. Next step of quasi-movement, as in Fig. 1b, is based on movement suppression by not exceeding an EMG detection threshold. EEG classification accuracy shown in this case, for some subjects, is almost perfect, with some sudden drops in accuracy due to possible EMG interference is noticeable. Finally the imagery-movement experiment presents accuracy improvement for all subjects in Fig. 1c. It is interesting to note, that the accuracy gain is obtained not only in *mu*-rhythm area.

The presented approach is a step forward in subjects training to perform movement imagery BCI/BMI experiments.

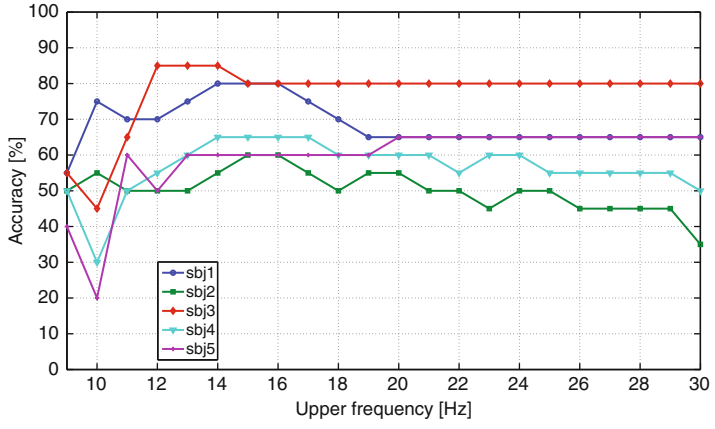
Acknowledgements The research presented in this paper was supported in part by the Japan Society for the Promotion of Science grant no. 21360179.



(a) Real-movement.



(b) Quasi-movement.



(c) Imagery-movement.

Fig. 1 The classification accuracy results

References

1. Pfurtscheller, G., Aranibar, A.: Event-related cortical desynchronization detected by power measurements of scalp EEG. Electroencephalogr. Clin. Neurophysiol. **42** (1977) 817–826
2. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain–computer interfaces for communication and control. Clin. Neurophysiol. **113** (2002) 767–791
3. Cichocki, A., Washizawa, Y., Rutkowski, T., Bakardjian, H., Phan, A.H., Choi, S., Lee, H., Zhao, Q., Zhang, L., Li, Y.: Noninvasive BCIs: Multiway signal-processing array decompositions. Computer **41**(10) (2008) 34–42
4. Wolpaw, J.R., McFarland, D.J.: Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proc. Nat. Acad. Sci. USA **101**(51) (December 2004) 17849–17854
5. Nikulin, V.V., Hohlefeld, F.U., Jacobs, A.M., Curio, G.: Quasi-movements: A novel motor-cognitive phenomenon. Neuropsychologia **46** (2008) 727–742
6. Niedermeyer, E., Da Silva, F.L., eds.: Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 5 edn. Philadelphia, PA: Lippincott Williams & Wilkins (2004)
7. Ramoser, H., Mueller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehab. Eng. Rehab. Eng **8**(4) (2000) 441–446

High Resolution Common Spatial Frequency Filters for Classifying Multi-class EEG

Qibin Zhao, Tomasz M. Rutkowski, Andrzej Cichocki, and Liqing Zhang

Abstract The Common Spatial Patterns (CSP) algorithm is a highly successful spatial filtering method for extracting spatial patterns related to specific mental tasks from electroencephalogram (EEG) signals. The performance of CSP highly depends on the selection of frequency band in the preprocess. However, the most discriminative frequency band features varies slightly with subjects and mental tasks. In order to provide high resolution in frequency domain, we propose an common spatial frequency patterns method to learn most discriminative spatial and frequency filters simultaneously for specific mental task. The results on EEG data during motor imagery (MI) tasks demonstrate the good performance of our method with decreased number of EEG channels.

Keywords EEG · BCI · CSP

1 Introduction

Recently, Brain Computer Interface (BCI), which aims to establish a direct communicate pathway between brain and computer, has been widely researched over the world [1]. To this end, CSP method [2, 3] has been proven to be very powerful in determining spatial filters which can extract discriminative brain rhythms. However the performance will suffer from the frequency band which is either selected manually or unspecifically set to a broad band filter. In [4, 5], several algorithms were presented to consider the frequency information during the optimization. But, the low frequency resolution and the difficulties of choosing the best parameters still remain a challenging problem. To provide higher frequency resolution, we propose Common Spatial Frequency Patterns (CSFP) which is able to find the spatial and frequency combination filters containing the maximum

Q. Zhao (✉)
Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, Saitama, Japan
e-mail: qbzhao@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_102,

discriminative information. Hence the performance outperforms traditional CSP on average, and in cases when the frequency range of most discriminative rhythm for mental tasks are different, a considerable improvement of classification accuracy can be achieved.

2 Method

Time-frequency analysis of event related potentials of EEG and MEG data has recently attracted increasing attention for BCI [6, 7]. To accommodate non-stationarity frequency analysis, continuous wavelet transform can obtain an improved trade off between temporal resolution and frequency resolution through varying the window length over frequencies.

The continuous wavelet transform (CWT) of the EEG signal $x^{c,k}(\tau)$ is defined as:

$$W^{c,k}(a,t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x^{c,k}(\tau)\psi(\frac{\tau - t}{a})d\tau, \tag{1}$$

where $W^{c,k}(a,t)$ represents the CWT of the data segment $x^{c,k}(\tau)$, c represents channel index and k represents EEG trial number. Although many types of wavelets exists, the Morlet wavelets are appropriate for time-frequency analysis of EEG signals [8]. Since the scale parameter a is related to frequency f by the relation $a = \omega_0/(2\pi f)$, we define

$$\mathbf{p}_k^{c,f} = W^{c,k}\left(\frac{F_c}{fT}, t\right), \tag{2}$$

where $\mathbf{p}_k^{c,f}$ represents the time-vary wavelet coefficients of k-th trial EEG at channel c and frequency f. Then, we can form a spatio-frequency distribution matrix for each channel and each frequency bin, i.e.,

$$\mathbf{Y}_k = \left(\mathbf{p}_k^{c_1,f_1}, \ldots, \mathbf{p}_k^{c_n,f_1}, \ldots, \mathbf{p}_k^{c_1,f_m}, \ldots, \mathbf{p}_k^{c_n,f_m}\right)^T \tag{3}$$

where $\mathbf{Y}_k \in R^{N\times M}$, N is the number of channels $\times$ frequency (i.e., n$\times$m) , M is the number of samples in each trial.

Based on this time-frequency representation, the class-covariance matrices are given as,

$$\Gamma^{(i)} = \sum_{k\in class_i} \frac{\mathbf{Y}_k\mathbf{Y}_k^T}{\mathrm{tr}(\mathbf{Y}_k\mathbf{Y}_k^T)}, \tag{4}$$

where $class_i$ refer to the i-th class of the training data. The optimization criterion of CSFP is to find maximal discriminative spatial and frequency filters described by

$\mathbf{w}_p \in R^d (d = n \times m)$ which maximizes the variance in one class while simultaneously minimizes the total variance in all class. In order to find filters for each class, we can set the two matrices $\mathbf{S}_I$ and $\mathbf{S}_T$ as follows:

$$\mathbf{S}_I = \Gamma^{(i)}|_{i=1}^{M},\ \mathbf{S}_T = \frac{1}{M}\sum_{i=1}^{M}\Gamma^{(i)}, \tag{5}$$

where $\mathbf{S}_I$ is class-specific covariance matrix and $\mathbf{S}_T$ is total covariance matrix of M class. Each filter $\mathbf{w}$ is therefore found by solving the following optimisation problem:

$$\underset{\mathbf{w}}{\text{argmax}}\ \frac{\mathbf{w}^T\mathbf{S}_I\mathbf{w}}{\mathbf{w}^T\mathbf{S}_T\mathbf{w}}. \tag{6}$$

The projection vector $\mathbf{w}$ can be be shown to be equivalent to solving the generalised eigenvalue problem $\mathbf{S}_I\mathbf{w} = \lambda\mathbf{S}_T\mathbf{w}$, where λ is a generalized eigenvalue, and $\mathbf{w}$ is a generalized eigenvector corresponding to λ.

By the optimizing criteria in Eq. (6), we can get projection matrix $\mathbf{W}^{(i)}$ i.e., spatial frequency filters, for each class i. The number of projection vector in $\mathbf{W}^{(i)}$ is selected by cross validation on training data. Finally, combining each $\mathbf{W}^{(i)}|_{i=1}^{M}$, we obtain $\mathbf{W}_{\text{csfp}} = [\mathbf{W}^{(1)}\mathbf{W}^{(2)}, \ldots, \mathbf{W}^{(M)}]$ which can be seen as spatial frequency filters for multi-class data. Therefore, the $\mathbf{Y}_k$ of new EEG trials can be projected onto $\mathbf{W}_{\text{csfp}}$ as:

$$\mathbf{Z}_k = \mathbf{W}^T\mathbf{Y}_k. \tag{7}$$

$\mathbf{Z}_k$ denotes the maximal discriminative components for multi-classes data.

3 Results and Discussions

For illustrations of the properties of the proposed method, we focus on classification of MI tasks, e.g. left hand, right hand and/or foot. The obvious variety of spatial distribution on different frequency band are observed in Fig. 1. To demonstrate the suitability of CSFP for small number of channels, only 5 channels (i.e., C3, Cp3, Cz, Cp4, C4) EEG were recorded from the scalp at a sampling rate of 256 Hz. Figure 2 demonstrates the variety of spectra among subjects and class-dependent spectra for

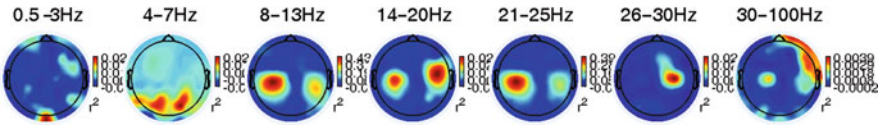


Fig. 1 Scalp topographies of r^2-values of left and right tasks on different frequency band

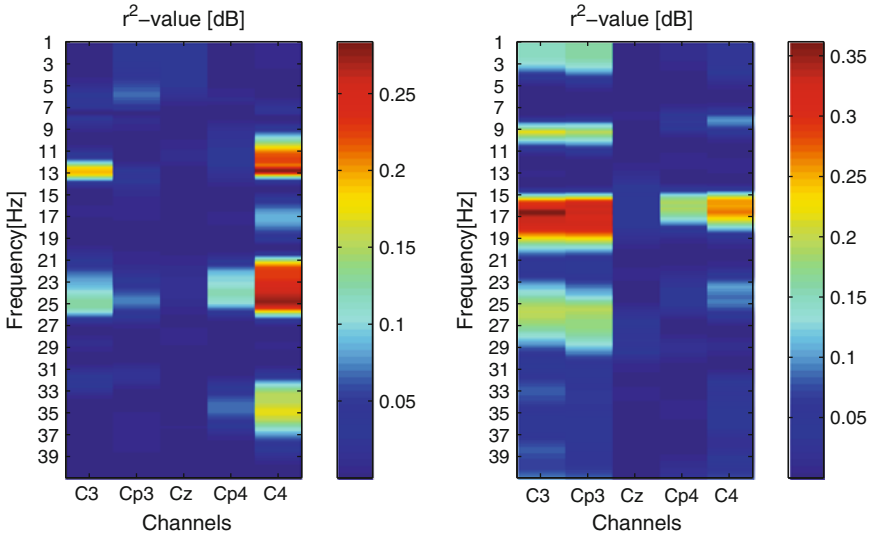


Fig. 2 r^2-values between left and right tasks for two subjects (i.e., *left figure* for S1, *right figure* for S2)

the same subject, such as the left class of S1 focuses on 11–13 Hz whereas right class focuses on 21–25 Hz. Therefore, the different weights on spectra are required evidently.

The spatial and frequency filters obtained by CSFP are showed in Fig. 3. Obviously, not only different spatial filters but also spectra filters for each class are learned from training data. The evident difference of two classes EEG can be found by both spatial and frequency patterns, which indicates that the frequency

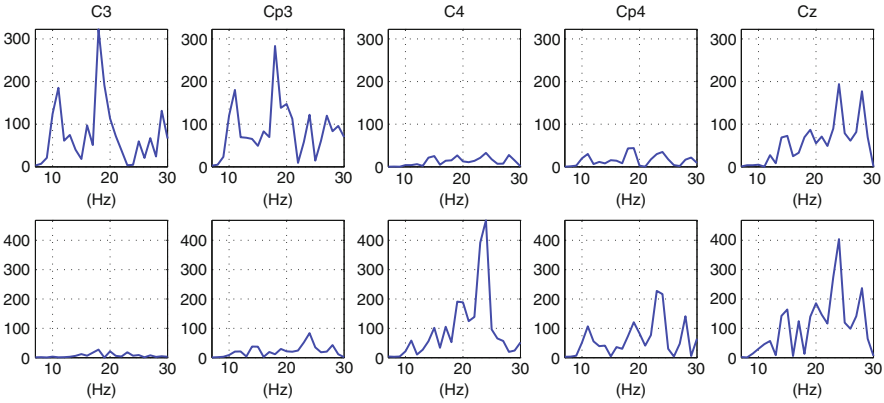


Fig. 3 Optimal spectra filters on each channels. *Upper row* are maximum filters for left class and *lower row* are for right class

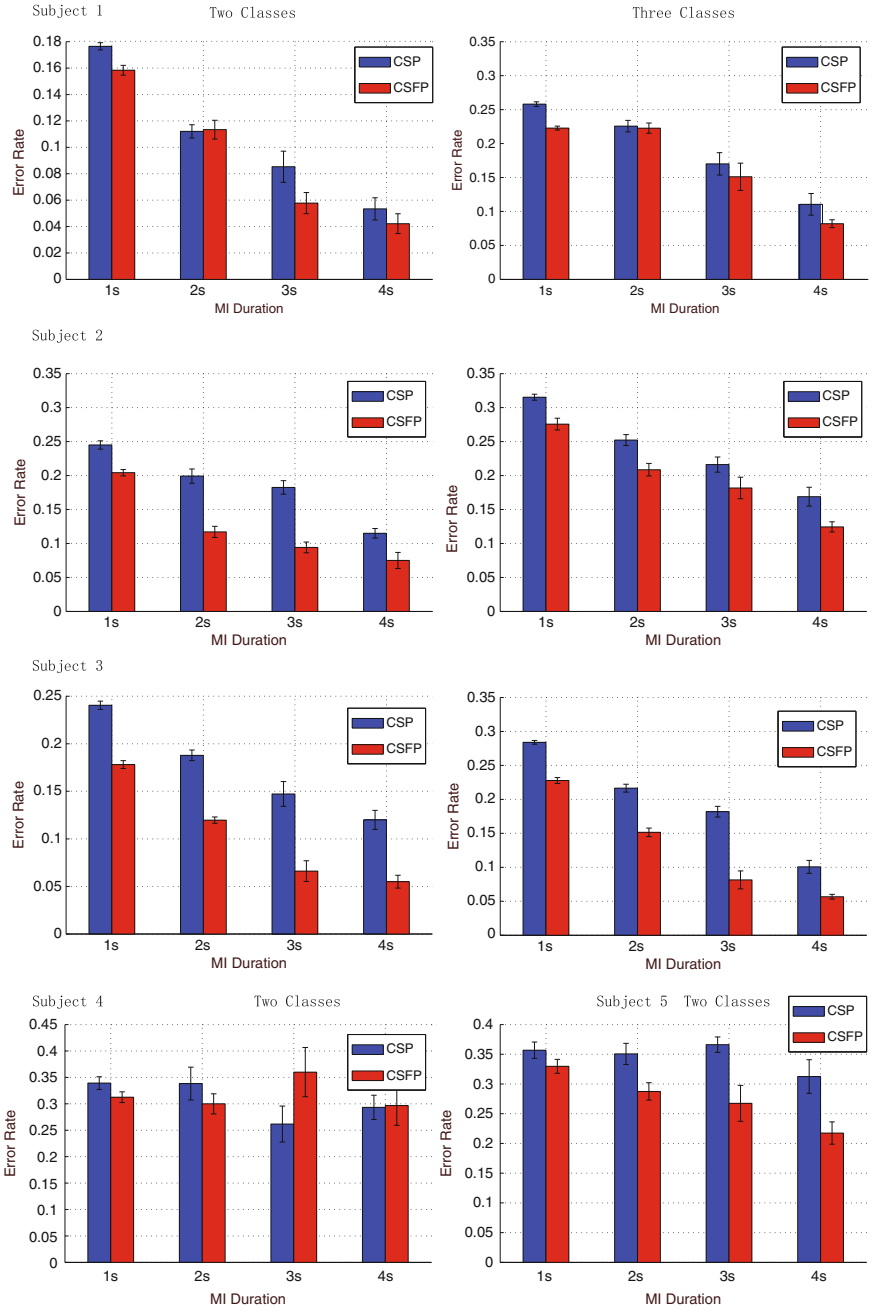


Fig. 4 Classification error rate and standard deviation of 10×10-fold cross validation for five subjects by CSP and CSFP with different MI duration. The two classes and three classes results are presented respectively. Three out of five subjects have the better performance whereas the other two subjects (S4 and S5) have the poor performance close to random results

information are not only subject-dependent but also class-dependent. As expected, there is a high discrimination in μ and β bands, but the low discrimination in other frequency band. So instead of having a spatial projection onto a broad band (8–30 Hz) signal as a solution, the CSFP yields an improved spatio-frequency resolution of the discriminative signals.

In order to illustrate the robustness of the proposed algorithm and provide the continuous EEG classification, the preprocess of segmentation has been applied to the original EEG signals. Each EEG trial has been segmented into several sub-trials by sliding window with fixed window length. After preprocessing, we can obtain overlapped EEG trials with duration of each trial varied between 1 to 4 s. By using linear Support Vector Machine (SVM) as classification model, the 10×10-fold cross validation error rate and standard deviation for 5 subjects are presented in Fig. 4. The results suggest that the proposed algorithm outperforms the CSP algorithm, in terms of an decreased classification error rate.

4 Conclusion

The presented CSFP algorithm successfully classifies EEG signals with fewer channels by optimizing the spatial and frequency patterns simultaneously. Due to the better results of CSFP on short MI duration, we can conclude that the CSFP feature extraction strategy will be more suitable for providing high performance and fast response for asynchronous BCI.

Acknowledgments The work was supported in part by the Science and Technology Commission of Shanghai Municipality (Grant No. 08511501701) and the National Natural Science Foundation of China (Grant No. 60775007).

References

1. Wolpaw, J.R., Birbaumer, N., et al.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113** (March 2002) 767–791.
2. Ramoser, H., Müller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. **8**(40) (Dec 2000) 441–446.
3. Müller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial eeg classification in a movement tast. Clin. Neurophysiol. **110** (1999) 787–798.
4. Lemm, S., Blankertz, B., Curio, G., Muller, K.: Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans. Biomed. Eng. **52**(9) (2005) 1541–1548.
5. Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G., Muller, K.: Optimizing spatio-temporal filters for improving brain-computer interfacing. Adv. Neural Inf. Process. Syst. **5** (2006) 315–322.
6. Bostanov, V., Kotchoubey, B.: The t-CWT: A new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics. Clin. Neurophysiol. **117**(12) (2006) 2627–2644.
7. Yamawaki, N., Wilke, C., Liu, Z., He, B.: An enhanced time-frequency-spatial approach for motor imagery classification. IEEE Trans. Neural Syst. Rehabil. Eng. (a publication of the IEEE Engineering in Medicine and Biology Society) **14**(2) (2006) 250.
8. Mørup, M., Hansen, L., Arnfred, S.: ERPWAVELAB A toolbox for multi-channel analysis of time–frequency transformed event related potentials. J. Neurosci. Methods **161**(2) (2007) 361–368.

Suitable ICA Algorithm for Extracting Saccade-Related EEG Signals

Arao Funase, Motoaki Mouri, Andrzej Cichocki, and Ichi Takumi

Abstract Our goal is to develop a novel BCI based on saccade-related EEG signals. It is necessary to analyze raw EEG signals in signal processing methods for BCI. In order to process raw EEG signals, we used independent component analysis (ICA). Previous paper presented extraction rate of saccade-related EEG signals by five ICA algorithms and eight window size. However, three ICA algorithms, the FastICA, the NG-FICA and the JADE algorithms, are based on 4th order statistic and AMUSE algorithm has an improved algorithm named the SOBI. Therefore, we must re-select ICA algorithms. In this paper, Firstly, we add new algorithms; the SOBI and the MILCA. Using the Fast ICA, the JADE, the AMUSE, the SOBI, and the MILCA, we extract saccade-related EEG signals and check extracting rates. Secondly, we check relationship between window sizes of EEG signals to be analyzed and extracting rates.

1 Introduction

Brain-computer interfaces (BCIs) have been researched as a novel human interface for a few decades. The capabilities of BCIs allow them to be used in situations unsuitable for the conventional interfaces. BCIs are used to connect a user and a computer via an electroencephalogram (EEG).

EEG related to saccadic eye movements have been studied by our group toward developing a BCI eye-tracking system [1]. In previous research, saccade-related EEG signals were analyzed using the ensemble averaging method. Ensemble averaging is not suitable for analyzing raw EEG data because the method needs many repetitive trials.

A. Funase (✉)
Graduate School of Engineering, Nagoya Institute of Technology, Nagoya 466-8555, Japan; Brain Science Institute, Wako, Saitama 351-0198, Japan
e-mail: funase.arao@nitech.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_103,

Recording EEG data repetitively is a critical problem to develop BCIs. It is essential to overcome this problem in order to realize practical use of BCIs for single trial EEG data.

Recently, the independent component analysis (ICA) method has been introduced in the field of bio-signal processing as a promising technique for separating independent sources. The ICA method can process raw EEG data and find features related to various one's activity. Therefore, the ICA algorithm overcomes the problems associated with ensemble averaging, and the ICA analyzes the waveforms of the EEG data.

There are many algorithms to compute independent components [2]. In previous studies [3], we used the FastICA, the NG-FICA, the AMUSE, the JADE to analyze saccade-related EEG signals. However, we must re-select an ICA algorithm since three ICA algorithms: the FastICA, the NG-FICA and the JADE algorithms are based on the 4th order statistic and the AMUSE algorithm has an improved algorithm named the SOBI [4].

In this research, we add new algorithms: the SOBI and the MILCA [5]. The SOBI is an improved algorithm based on the AMUSE and uses the independency based on two covariance matrices at different time steps. The MILCA uses the independency based on mutual information. Using the Fast ICA, the JADE, the AMUSE, the SOBI, and the MILCA, we extract saccade-related EEG signals and check extracting rates.

Secondly, we focus on window sizes of EEG signals to be analyzed. In order to analyze EEG signals in on-line system, we must choose an appropriate window size to extract continuous EEG signals. In this paper, we separate window sizes into two groups: the windows excluding EEG signals after eye movements and the windows include EEG signals after eye movements.

2 Experimental Settings

In order to record saccade-related EEG signals, we performed visually guided saccade task. This task was to record the EEG signals during a saccade to a visual target that is either his/her right side or left side. This experiment was comprised of 50 trials in total: 25 on the right side and 25 on the left side. The number of subjects is 5. Their age is from 22–24 years old.

The EEG signals were recorded through 19 electrodes (Ag–AgCl), which were placed on the subject's head in accord with the international 10–20 electrode position system. The Electrooculogram (EOG) signals were simultaneously recorded through two pairs of electrodes (Ag–AgCl) attached to the top–bottom side and right-left side of the right eye.

Recorded EEG signals were calculated by five ICA algorithms: FastICA, AMUSE, JADE, SOBI, MILCA. In order to calculate independent components, we must decide the window length. In this paper, there were 8 size windows.

Group 1

1. Window A: –999–1000 ms
2. Window B: –499–500 ms
3. Window C: –349–350 ms

Group 2

4. Window D: –999–0 ms
5. Window E: –499–0 ms
6. Window F: –349–0 ms
7. Window G: –249–0 ms
8. Window H: –99–0 ms

0 [ms] indicates the starting point of saccade. After eye movements, EEG signals include big noises caused by EOG signals. In order to observe influence of noises caused by EOG signals, we separated window size into two groups: Window A to C including EEG signals after saccade and window D to H excluding EEG signals after saccade.

In using five algorithms and eight windows, we calculated saccade-related independent components.

3 Experimental Results

First, we define two words: an extracting rate and saccade-related IC. The extraction rate is defined by the following ratio:

(*the number of trials in which saccade-related IC are extracted*)/(*The total number of trials*)

We make assumption that a saccade-related IC has a positive peak from –50 ms ∼ –1 ms. The peak-amplitude n is larger than 3; $n = \frac{\bar{x}-\mu}{s}$; where $\bar{x}$ is mean of EEG potential during 1,000 ms before saccade, μ is maximum amplitude, and s is standard deviation during 1,000 ms before saccade.

Table 1 represents the rate for extracting saccade-related ICs from the raw EEG data by each algorithm in the case of window E. From these results, the FastICA and JADE got good performance in extracting saccade-related independent components. However, the results of the AMUSE and SOBI and MILCA algorithm were not good. From these results, in order to extract saccade-related EEG signals, it is not suitable to use independency of 2nd order statistics and the mutual information.

Next, we focus on extracting rate in each windows (see Table 2). From Table 2, extracting rates in group 1 were lower than those in group 2. Therefore, we should

Table 1 Extracted rate by four ICA algorithms

	AMUSE(%)	FICA(%)	JADE(%)	SOBI(%)	MILCA(%)
A	14	98	100	70	50
B	18	82	94	76	46
C	30	94	96	80	62
D	30	98	98	66	50
E	24	94	96	70	46

Table 2 Extracted rate by eight window size

Group	Window size(ms)	FastICA(%)	JADE(%)
1	–999 ∼ 1000	37.2	38
	–499 ∼ 500	29.6	27.2
	–349 ∼ 350	22.4	26.4
2	–999 ∼ 0	90	93.6
	–499 ∼ 0	93.2	96.4
	–349 ∼ 0	99.4	99.2
	–249 ∼ 0	93.2	93.6
	–99 ∼ 0	99.4	99.2

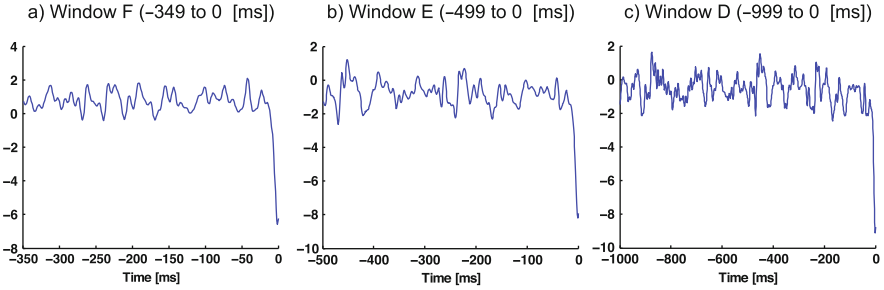


Fig. 1 Extracted signals for FastICA by each window size

not use EEG signals after saccade. because the signals in group 1 include EOG noise. In the case of group 2, the results of small window size is better. From these result, we can get good results in the case of short window size excluding signals after saccade.

Figure 1 is extracted signals by FastICA and window D,E,F. Each extracted signals are not the same although input signals are the same. However, each signals denote the same tendency.

4 Conclusion

This paper presented extraction rates of saccade-related EEG signals by five ICA algorithms and eight window sizes.

As results of extracting rate focused on ICA algorithms, The JADE and Fast ICA had good results.

As results of extracting rates focused on window sizes, the window H (–99 ∼ 0 ms) had good results. In the case of the window A, B, and C, we could not get good results because these windows included big EOG noise.

In next step, we must check relationship between extracting rate and the number of input channels. In order to develop BCI, we must select a few input channels instead of present input channels; 19 channels.

References

1. Funase, A., Yagi, T., Kuno, Y., Uchikawa, Y.: A study on electro-encephalo-gram (EEG) in eye movement. Studies in Appl. Electromagn. Mech. **18** (2000) 709–712.
2. Cichocki, A., Amari, S.: Adaptive Blind Signal and Image Processing. New York, NY: Wiley (2002).
3. Funase, A., Hashimoto, T., Yagi, T., Barros, A.K., Cichocki, A., Takumi, I.: Research for estimating direction of saccadic eye movements by single trial processing. Proceedings of 29th Annual International Conference of the IEEE EMBS (2007) 4723–4726.
4. Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., Moulines, E.: Second-order blind separation of temporally correlated sources. Proceedings of the International Conference on Digital Siginal Processing (Cyprus) (1993) 346–351.
5. Stogbauer, H., Kraskov, A., Astakhov, S.A., Grassberger, P.: Least dependent component analysis based on mutual information. Phys. Rev. E **70**(6) (2004) 066123.

EEG Based Brain-Computer Interface System for Remote Vehicle Controlling

Han Sun, Liqing Zhang, Jie Li, and Qibing Zhao

Abstract Brain Computer Interface (BCI) bridges human brain and computer. In this paper, we propose a real-time BCI system that enables subject to control remote vehicle by motor imagery. Multilinear generation of common spatial pattern analysis is introduced to extract high-order tensor EEG features. The general framework of BCI interfacing paradigm for remote vehicle control is then developed. To validate the performance of the developed model, both offline simulation and real-time experiment are given. The real-time experiment shows after a proper period of training, subjects can control the vehicle running along a designated path smoothly.

1 Introduction

As a newly rising computational neuroscience and cognitive neurodynamics research direction, brain computer interface technology has encouraged incremental attentions [1]. While BCI technology originally aimed to establish a direct communication channel between brain and computer for paralyzed people, recent BCI research has expanded to a broader area that may benefits normal people as well, for instances, remote system controlling, virtual environment exploring, and new generation game controlling. Meanwhile, among available techniques for brain signal acquisition, EEG is preferred for real time BCI systems due to its high temporal resolution, portable size, noninvasive recording and lower cost [2].

Recent studies have discovered human brain can adapt to BCIs that predict motor imagery EEG patterns and control external devices [3]. In this paper, we use multilinear generation of common spatial pattern to analyze single trial EEG in motor (left and right hand) imagery task, and implement a real-time BCI system that enable subject to control remote vehicle.

H. Sun (✉)
Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: sunhan.cn@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_104,

2 Methods

In the signal preprocessing module, we used constrained fast ICA to remove eye movement artifacts. Before experiment, subjects were required to move the eyeball of purpose, during which the EOG contribution to the other EEG channels was approached. Then after restoring the signals by deflation procedure, the artifacts-corrected EEG signals were obtained.

Common Tensor Discriminant Analysis(CTDA) [4] was employed in the feature extraction module. The idea of CTDA is to find the most discriminative tensor subspace by building $K-1$ filters W_i, $i = 1...M-1$, where different classes have the most differing power ratios after being projected. And the M-order tensor data X is projected as follow:

$$Z = X \prod_{d=1}^{M-1} \times_d W_d^T$$

Then the feature vector F is calculated:

$$F = log\left(\frac{diag(Z_M Z_M^T)}{tr(Z_M Z_M^T)}\right)$$

Z_M is the matrix form of tensor Z where the Mth dimension maps to the rows of the matrix. In this research, we transformed the original data to *channel* $\times$ *frequency* $\times$ *time* tensor, where $M = 3$.

3 System Design and Experiments

In this section, we will briefly introduce the integrated system and BCI interfacing paradigm for remote vehicle control.

3.1 System Architecture

Here we propose a general framework and development platform for developing an efficient, reliable and extendible BCI system.

Shown in Fig. 1, the entire system is composed of several subsystems and modules. The console mainly controls all functional processes. After configuring the parameters and options, it begins to process EEG signals and translate the subject's intention to internal commands. Synchronously, signals and data are transmitted to visualization system, where data are illustrated for experiment operator. Finally, internal commands are packaged and delivered to control remote vehicle throughout a transmission interface.

Functionally, our system provides two working modes. Before a subject begins real-time experiment, several runs of training procedures are taken in training mode, while a statistical model is trained for the specific subject. Another one is online

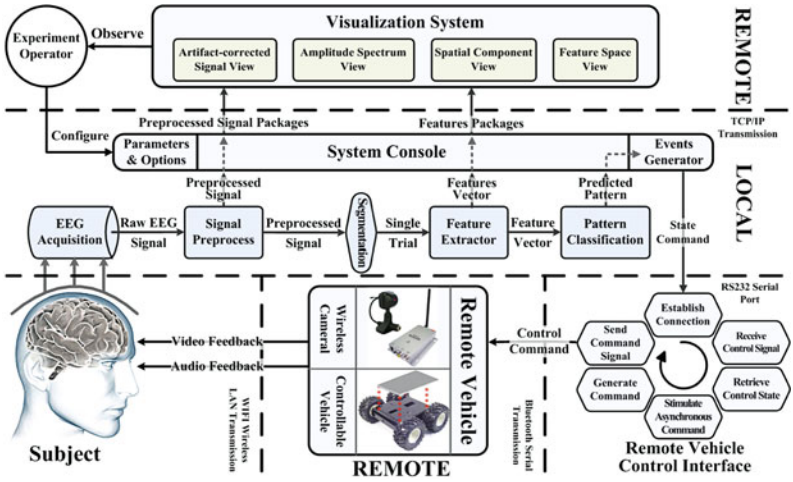


Fig. 1 BCI system framework diagram

processing and real-time control mode, in which EEG signals are processed and translated to remote control commands with real-time capability.

3.2 Experiments and Data Acquisition

To validate our BCI system, we carried out an experiment with four right-handed healthy subjects. For each subject, the experiment included a training procedure consisted of 6 sessions of 60 trials each and an online experiment. Focusing on the screen where the image captured by a camera on the vehicle was shown in first person view, subject was required to drive along a designated path.

Twenty one channels of EEG and 2 channels of EOG signals were recorded and sampled at 250 Hz.

4 Results and Analysis

In the training stage, four subject's EEG signals were recorded and analyzed. We use different trail lengths to approach the best model for specific subject. The classification accuracies are illustrated in Table 1. Three of the subjects' best accuracies

Table 1 Classification accuracy for each subject using different trail lengths

Subject	1.0 s	1.5 s	2.0 s
Sub1	86.1%	93.0%	**98.3%**
Sub2	**78.8%**	73.0%	75.0%
Sub3	87.6%	92.5%	**95.8%**
Sub4	92.0%	90.4%	**95.7%**

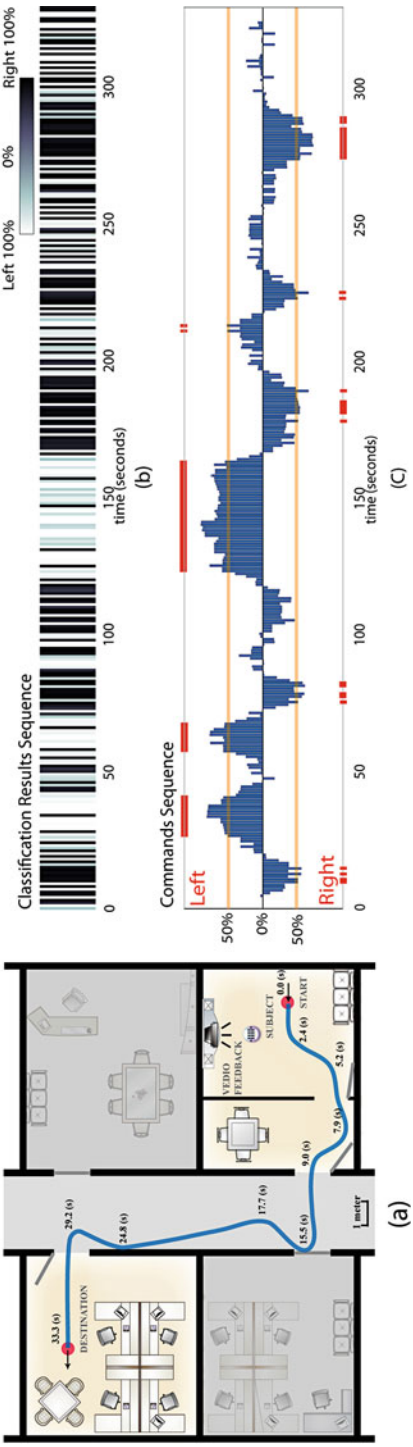


Fig. 2 Real time experiment result. (**a**) Map of pathway that the vehicle was steered, in which it started from one room, went through the corridor, and arrived at the destination in another room within 33 s. (**b**) The classification outputs of controlling indicated by predict probability. (**c**) The command sequence (*red dots*), which was obtained by averaging the classification outputs (*blue bars*) within the latest 1 s

were greater than 95%, which inferred the EEG patterns for different mental tasks were predictable for most subjects.

After a proper time of self-adaption and training, three subjects were able to control the remote vehicle smoothly. Figure 2 illustrate the process of subject1's online experiment. He was required to drive the vehicle from one point to another. Such real navigation could have favorable impact on EEG dynamics, making the whole procedure easier for general subjects.

5 Discussion

To sum up, the developed BCI system was successfully applied to control remote vehicle. However the task of developing more robust and universal applicable BCI device that could be utilized in daily life remains changeling. The main future work may focus on exploring more efficiency feature selection methods, discovering new user interaction scheme and introducing advanced control theory to strengthen the maneuverability of target device.

Acknowledgments The work was supported by the National Basic Research Program of China (Grant No. 2005CB724301) and the Science and Technology Commission of Shanghai Municipality (Grant No. 08511501701).

References

1. Wolpaw, J.R., Birbaumer, N., McFarland, D.J.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**(8) (2002) 767–791.
2. Pfurtscheller, G., Neuper, C.: Motor imagery activates primary sensorimotor area in humans. Neurosci. Lett. **239** (1997) 65C68.
3. Zhao, Q.B., Zhang, L.Q., Cichocki, A.: EEG-based asynchronous BCI control of a car in 3D virtual reality environments. Chinese Sci. Bull. **54** (2009) 78–87.
4. Zhao, Q.B., Zhang, L.Q., Cichocki, A.: Multilinear generation of common spatial pattern. International Conference on Acoustic, Speech, and Signal Processing (2009)

Channel Selection for Motor Imagery-Based BCIs: A Semi-supervised SVM Algorithm

Jinyi Long, Yuanqing Li, and Zhenghui Gu

Abstract Given a frequency band, we propose a semi-supervised SVM algorithm to select a set of channels for motor imagery-based BCIs in this paper. Both training data with labels and test data without labels are used for training a classifier in this approach. Hence, it is suitable for the case of small training data set. To test our algorithm, it is applied to a BCI competition data set. Data analysis results demonstrate the effectiveness of our algorithm.

Keywords Electroencephalogram (EEG) · Motor imagery · Brain computer interface (BCI) · Channels · Semi-supervised learning

1 Introduction

For a motor imagery-based BCI system, channels and frequency band are two important parameters for effecting the performance of a BCI system. This paper discusses channel selection for motor-imagery based BCI systems. Generally, channels are manually set or selected through cross-validation. The state-of-the-art approaches for rating the relevance of channels include zero-norm optimization based-support vector machine (SVM), Recursive Feature Elimination (RFE) and Fisher ratio [1]. After ranking all channels, the number of selected channels was generally determined by the average classification accuracy obtained in cross validation of training data set [1, 2].

The above methods for parameters selection generally need sufficient training data to obtain reliable channels. However, collecting sufficient training data is often boring and time consuming. For channel selection, we extend an iterative semi-supervised SVM algorithm mentioned in [3, 4], where it is proposed for joint feature re-extraction and classification considering small training data. In our algorithm,

Y. Li (✉)
College of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China
e-mail: auyqli@scut.edu.cn

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_105,

Fisher ratio is applied to rank channels, while Rayleigh coefficient is applied to determine the number of channels. The computation of Fisher ratio and Rayleigh coefficient is based on both training data and test data.

To validate the proposed algorithm, we apply it to the data set IVa of BCI competition III in 2005 [5]. Furthermore, we compare our algorithm with an existing method (Fisher ratio based cross-validation) mentioned in [1] for channel selection. Our data analysis results show that our algorithm can effectively perform channel selection with small training data set and provide satisfactory classification accuracy.

2 A Semi-supervised SVM Algorithm for Channel Selection

Given two raw EEG data set: A small training data set D_c containing N_c trials of EEG signals with labels $y_i \in \{+1, -1\}, (i = 1, \ldots, N_c)$ and a test data set D_t containing N_t trials of EEG signals. In the following, we present the outline of an iterative semi-supervised SVM algorithm for channel selection.

Algorithm 1

Step 1 For the first iteration, do the following Steps 1.1–1.4:

1.1 (Feature extraction and classification) Given a frequency band F_0, band-pass filter is applied to the raw EEG data set $D = D_c \cup D_t$. We train a spatial transformation matrix with training data set D_c to extract the CSP features. The extracted feature vectors are denoted as $\mathbf{x}_k, k = 1, \ldots, N_c, N_c + 1, \ldots, N_c + N_t$. With the feature vectors and the known labels corresponding to training data D_c, we train a SVM model and predict the class labels $\{y_{N_c+1}, \ldots, y_{N_c+N_t}\}$ for the test data set D_t.

1.2 (Channel selection) Using the data set$D$ with the given labels and the predicted labels, we calculate Fisher ratio $FR_l^{(1)}$ for the lth $(l = 1, \ldots, n)$ channel and rank all channels using $FR_l^{(1)}$. Define $n - 1$ subsets of channels $Q_j^{(1)}$ $(j = 2, \ldots, n)$ which contain j channels with the highest Fisher ratio scores and then calculate the Rayleigh coefficient $RC_j^{(1)}$ for the subset $Q_j^{(1)}$. A channel subset denoted as $Q^{(1)}$ is then determined by the maximal difference between $Q_j^{(1)}$ and $Q_{j+1}^{(1)}$. In this step, the superscript 1 represents the first iteration.

1.3 (Data updation) Using the selected channels $Q^{(1)}$, we construct data set $D_c^{(1)}$ which contains N_c trials of raw EEG signals from D_c. We also construct data set $D_t^{(1)}$ which contains N_t trials of raw EEG signals from D_t. Denote $D^{(1)} = D_c^{(1)} \cup D_t^{(1)}$.

1.4 (Feature re-extraction and classification) Based on the data set $D^{(1)}$, we re-extract the CSP feature vectors $\{\mathbf{x}_1^{(1)}, \ldots, \mathbf{x}_{N_c+N_t}^{(1)}\}$ as Step 1.1, then train a SVM model and perform classification. The predicted labels are denoted as $y_{N_c+1}^{(1)}, \ldots, y_{N_c+N_t}^{(1)}$ for test data set $D_t^{(1)}$.

Step 2 For the kth iteration ($k = 2, 3, \ldots$), do the following Steps 2.1–2.4.

2.1 (Channel selection) Based on the data set D with the labels for D_c and the labels predicted in the $(k-1)$th iteration for D_t in which the labels is the same as $D_t^{(k-1)}$, we perform channel selection as in Step 1.2. The subset of selected channels is denoted as $Q^{(k)}$.

2.2 (Data updation) We construct data set $D_c^{(k)}$ and data set $D_t^{(k)}$ similar to Step 1.3 with $Q^{(k)}$. Denote $D^{(k)} = D_c^{(k)} \cup D_t^{(k)}$.

2.3 (Feature re-extraction and classification) This step is similar to Step 1.4 and the predicted labels are denoted as $y_{N_c+1}^{(k)}, \ldots, y_{N_c+N_t}^{(k)}$.

2.4 (Calculation of termination index) Calculate a ratio which reflects the difference between the labels predicted in the kth iteration and the labels predicted in $(k-1)$th iteration:

$$r(k) = \frac{\sum_{i=N_c+1}^{N_c+N_t} |y_i^{(k)} - y_i^{(k-1)}|}{2N_t} \tag{1}$$

Step 3 (Termination) For a predefined constant δ_0, if $r(k) < \delta_0$, the algorithm terminates after k iterations. $Q^{(k)}$ is the set of selected channels, while $y_{N_c+1}^{(k)}, \ldots, y_{N_c+N_t}^{(k)}$ are the predicted labels for test data set D_t.

3 Data Analysis

In this section, we apply Algorithm 1 to data set IVa provided by the BCI competition III to demonstrate its effectiveness [5]. The data set for each subject is partitioned into two subsets: the first one containing the first 200 trials is used for 8-fold cross-validation and the remainder is used as an independent test data set which is not involved in the training process. One fold of EEG signals is used for training, while the other folds of EEG signals are used as the test data set. In this paper, we present our analysis results for the first 3 subjects "aa", "al" and "aw", as the results for the other two subjects are similar.

Data preprocessing steps include spatial filtering (Common Average Reference, CAR), band pass filtering in the range of 8–14 Hz corresponding to mu rhythm. Since ERD generally happens in sensorimotor cortex, we only use 60 channels located in the center area.

For each subject, we obtain two average accuracy curves over 8-fold cross validation for the test data set and the independent test set respectively. The three subplots of Fig. 1 show the iteration curves of averaged classification accuracies obtained by Algorithm 1 (solid lines) for the first three subjects respectively. From these subplots, we can see that most of these curves of prediction accuracies obtained by Algorithm 1 (solid lines) show increasing tendency and Algorithm 1 generally converges in about 10 iterations.

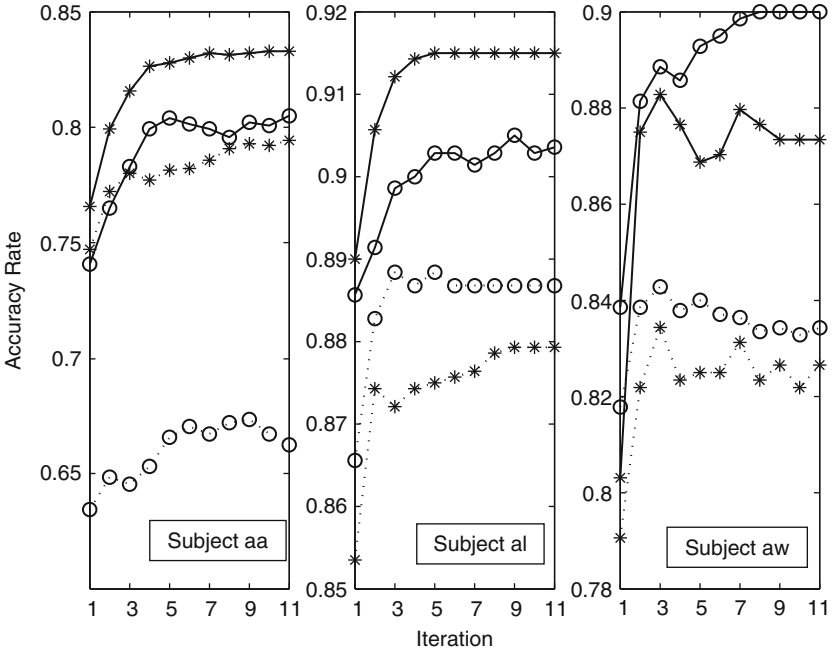


Fig. 1 Curves of averaged prediction accuracies for the first three subjects. (The *curves* with "*o*" correspond to test data sets used in 8-fold cross-validation, while the the curves with "∗" correspond to independent test sets. The *real line* is for Algorithm 1, while the *dashed line* is for comparison with Fisher ratio.)

For comparison, we now modify Algorithm 1 as follows. The channels are ranked by Fisher ratio and the number of selected channels is determined by cross validation (leave-one-out) on the small training data set. When the subset of channels is determined, an iterative semi-supervised SVM algorithm by removing all steps related to channel selection in Algorithm 1 is used for joint feature extraction and classification. Now we apply this approach to the data set and perform similar data analysis as above. Figure 1 shows the iterative curves of averaged classification accuracies (dashed lines) for the first three subjects.

Through Fig. 1, we can see that the performance of Algorithm 1 is better than that of the modified method. The main reason is that channel selection of the modified method is based on the small training data set, while the channel selection of Algorithm 1 is performed by a semi-supervised approach in which both the small training data set and the test data set are involved.

4 Conclusions

In this paper, we presented an iterative semi-supervised SVM algorithm for channel selection which is designed for motor-imagery based BCI. Since this algorithm

is based on semi-supervised learning strategy, it is suitable for the case of small training data set. Through analyzing the data set IVa in BCI competition III, the effectiveness of our algorithms were demonstrated.

Acknowledgements This work was supported by National Natural Science Foundation of China under Grant 60825306.

References

1. Lal, T., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., Scholkopf, B.: Support vector channel selection in BCI. IEEE Trans. Biomed. Eng. **51**(6) (2004)1003–1010.
2. Al-Ani, A., Al-Sukker, A.: Effect of feature and channel selection on EEG classification. Proceedings of the International Conference of IEEE EMBS (2006) 2171–2174.
3. Li, Y., Guan, C.: Joint feature re-extraction and classification using an iterative semi-supervised support vector machine algorithm. Mach. Learn. **71**(1) (2008) 33–53.
4. Li, Y., Guan, C., Li, H., Chin, Z.: A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. Pattern Recognit Lett **29**(9) (2008) 1285–1294.
5. Dornhege, G., Blankertz, B., Curio, G., Muller, K.: Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. IEEE Trans. Biomed. Eng. **51**(6) (2004) 993–1002.

Part XII
Neural Dynamics of Brain Disorders

On the Early Diagnosis of Alzheimer's Disease from EEG Signals: A Mini-Review

Justin Dauwels, François-Benoît Vialatte, and Andrzej Cichocki

Abstract In recent years, various computational approaches have been proposed to diagnose Alzheimer's disease (AD) from EEG recordings. In this paper, we review some of those approaches, and discuss their limitations and potential.

1 Introduction

Alzheimer's disease (AD) is a common neurodegenerative disorder that affects more than 10% of Americans over age 65, and nearly 50% of people older than 85; it is estimated that the prevalence of the disease will triple within the next 50 years [1]. No cure for Alzheimer's disease has been developed yet, but a number of medications are believed to delay the symptoms (and perhaps causes) of the disease. The progression of the disease can be categorized in four different stages. The first stage is known as Mild Cognitive Impairment (MCI), and corresponds to a variety of symptoms (most commonly amnesia) which do not significantly alter daily life. Between 6 and 25% of people affected with MCI progress to AD every year. The next stages of Alzheimer's disease (Mild and Moderate AD) are characterized by increasing cognitive deficits, and decreasing independence, culminating in the patient's complete dependence on caregivers and a complete deterioration of personality (Severe AD) [2].

Early diagnosis of Alzheimer's disease, and in particular diagnosis of MCI and Mild AD, is important, since medications are most effective at an early stage of the disease. Diagnosis of Alzheimer's disease is hard, however, and symptoms are often dismissed as normal consequences of aging. Diagnosis is usually performed through a combination of extensive testing and eliminations of other possible causes. Psychological tests such as Mini Mental State Examinations (MMSE), blood tests,

J. Dauwels (✉)
Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology (MIT), Cambridge, MA, USA
e-mail: justin@dauwels.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_106,

spinal fluid, neurological examination, and increasingly, imaging techniques are used to help diagnose the disease [3].

In the last years, several research groups have started investigating the potential of electroencephalograms (EEGs) for diagnosing AD. Since EEG recording systems are inexpensive and (potentially) mobile, EEG may potentially be used as a tool to screen a large population for the risk of AD.

In this paper, we review several state-of-the-art signal processing methods to detect perturbations in EEG signals caused by AD (Section 2). We assess the potential and limitation of such computational approaches (Section 3). At the end of the paper, we briefly address future challenges and open problems (Section 4). This paper is probably one of the few studies in recent years to provide an overview and critical assessment of various state-of-the-art signal processing methods for diagnosing AD from EEG signals.

2 Analysis of EEG of MCI and AD Patients

Studies have shown that AD has (at least) three major effects on EEG (see [4] for an in-depth review): slowing of the EEG, reduced complexity of the EEG signals, and perturbations in EEG synchrony. Those effects, however, are not always easily detectable: there tends to be a large variability among AD patients. As a result, none of those phenomena allow at present to reliably diagnose AD at an early stage. Many recent studies have investigated how to improve the sensitivity of EEG for detecting AD. In the following we briefly review some of these studies; we will treat the three major effects of AD on EEG separately.

2.1 Slowing of EEG

Many studies have shown that Alzheimer's disease (AD) causes EEG signals to slow down (see, e.g., [5–14]): AD is associated with an increase of power in low frequencies (delta and theta band, 0.5–8 Hz) and a decrease of power in higher frequencies (alpha and beta, 8–30 Hz, and gamma, 30–100 Hz). To quantify the changes in spectral power, one has applied Fourier transforms [5–11, 14] and sparsified time-frequency maps ("bump models") [12, 13].

2.2 Reduced Complexity of EEG Signals

The EEG of MCI and AD patients seems to be more regular than of age-matched control subjects [15–17]. The following complexity measures have been used to quantify this reduction in EEG complexity: approximate entropy [15], auto mutual information [18], sample entropy [15, 17], multiscale entropy [15], Lempel-Ziv complexity [15], and fractal dimension [16].

2.3 Perturbations in EEG Synchrony

Numerous studies have reported decreased EEG synchrony in MCI and AD patients under resting conditions ("spontaneous EEG") (see, e.g., [8, 18–32]). More precisely, the statistical dependence between spontaneous EEG signals recorded from different channels seems to be generally lower in MCI and AD patients than in age-matched control subjects. A large variety of measures have been applied to quantify this loss in statistical dependence, e.g., Pearson correlation coefficient [25], coherence [25, 26, 28], Granger causality [18, 19, 25], information-theoretic [32, 25] and state space based synchrony measures [8, 20, 25, 29, 30], phase synchrony indices [8, 25, 27, 28, 31], and stochastic event synchrony [21–25]. In a recent study, the spatial distribution of EEG phase synchrony in AD patients has been investigated [33]. By means of closely-related graph-theoretic methods, several studies have shown that EEG signals of MCI and AD patients have weaker small-world network characteristics compared to age-matched control subjects [34, 35].

The observed loss in synchrony is often attributed to a functional disconnection of the neocortex; it is probably not simply due to a loss of cortical neurons. For example, it may result from anatomical disconnections among different cortical regions in combination with reduced cholinergic coupling between cortical neurons [32]. In particular, a common hypothesis is that basal forebrain neurons may be severely affected in AD, resulting in a cerebral cholinergic deficit that leads to memory loss and other cognitive symptoms [32].

Interestingly, in a few studies that investigate the EEG of MCI and AD patients recorded during working memory tasks, an *increase* of EEG synchrony was observed in those patients [29, 36]. This inverse effect is often interpreted as the result of a compensatory mechanism in the brain.

One should keep in mind, however, that it is hard to directly interpret results obtained with synchrony measures. The synchrony of EEG signals may be significantly affected by brain events other than changes of synchrony, and by choices (like the reference electrode) that necessary have to be made during the analysis. Furthermore, as a single active source in the brain may affect the EEG signals across the entire scalp, changes in synchrony, and especially simultaneity of some events across channels, may be observed when the activity of one source alone changes, which is remote from a change in synchrony. As an alternative, one may use inversion methods to reconstruct sources, and next apply synchrony measures to those sources. However, the inversion problem is known to be notoriously difficult, and any source reconstruction method relies on certain assumptions, which may not necessarily hold.

3 Discussion

As pointed out earlier, EEG seems an attractive brain imaging modality for diagnosing AD, since EEG recording systems are inexpensive and (potentially) mobile. Moreover, in contrast to most other non-invasive brain imaging methods, EEG

has high temporal resolution, and may therefore contain crucial information about abnormal brain dynamics in AD patients.

Numerous studies have investigated the potential of EEG as a diagnostic tool for AD. At present, however, it is hard to assess whether EEG is truly useful for diagnosing AD. First of all, most studies report the results of statistical tests (p-values) without statistical post-correction. Since typically one conducts multiple tests simultaneously (sometimes hundreds or even thousands), e.g., individual pairs of electrodes or frequency bands, it is important to eliminate false positives. To this end, one may apply Bonferroni post-correction, or more powerful alternatives, such the false-discovery-rate correction method of Storey [37]. Unfortunately, not all studies on diagnosing AD from EEG signals apply such post-correction methods, and therefore, it is not always obvious how to interpret the reported results.

Second, few studies conduct discriminant analysis (linear or non-linear discriminant analysis, using support vector machines, neural networks, etc.); studies that do conduct discriminant analysis typically only report results for training data. The reported results are therefore often prone to overfitting, and they may be overoptimistic. To obtain more reliable classification results, one may for example apply crossvalidation, as has indeed been done in a handful studies (e.g., [6, 17, 25]).

Third, in most existing studies, a single measure to detect EEG abnormalities is applied to a single EEG data set. Since almost every study considers a different measure and a different EEG data set, it is hard to compare existing studies and to verify whether results are consistent.

Fourth, it is likely that one will need to combine various EEG characteristics in order to obtain a good diagnostic tool for AD, e.g., based on slowing, loss in EEG synchrony and complexity, and other features yet to be discovered. However, few studies systematically investigate large collections of EEG features (e.g., [6, 25]); it would be of great interest to apply dimensionality reduction methods to hundreds or even thousands of EEG features, to determine the most discriminative EEG features in a disciplined and statistically rigorous fashion. Moreover, it still needs to be verified whether the effects listed in Sections 2.1, 2.2, and 2.3 are independent. For example, it may be that EEG slowing and loss of EEG complexity are two sides of the same coin.

4 Conclusions

To conclude this paper, we point out several remaining challenges and topics for future research. At present, it is fairly difficult to gain access to EEG data of MCI or AD patients. Such databases are not publicly available, in contrast to ECG and other biomedical data (e.g., [38]). As a result, it is hard to systematically benchmark and assess the existing methods for diagnosing AD from EEG signals. Moreover, virtually none of those methods incorporate biophysical knowledge about AD (but see [39]); detailed mathematical models of the pathology of AD, in conjunction with EEG data analysis, may help us to improve the diagnosis of AD. Along the

same lines, one may expect further improvement by combining EEG with other imaging modalities, such as MRI (see, e.g., [40, 41, 18]), dMRI [42], TMS [43], and SPECT [44].

The correlation between AD risk factors (e.g., high plasma concentration of homocysteine [45]) and EEG characteristics needs to be investigated in greater detail. In addition, at present, the precise relation between the decline of cognition and memory and EEG abnormalities in AD patients remains largely unexplored (but see [46, 7, 47, 31]). It is also of great importance to investigate whether EEG helps to distinguish between MCI and different stages of AD (see, e.g., [48]), and between AD and other dementias (see, e.g., [49–53]).

An important degree of freedom is the EEG recording condition: one may record EEG: (i) while the subject is at rest (with open or closed eyes); (ii) while the subject performs working-memory or other tasks; (iii) while the subject is being stimulated with auditory, visual, tactile, or other signals (see, e.g., [54–58]). Depending on the recording situation, EEG signals may be more or less discriminative for MCI and AD; a systematic exploration of different recordings conditions with the aim of diagnosing MCI and AD needs to be conducted.

One may also analyze the EEG of animal models of AD (see, e.g., [59]), where the progression of AD can be rigorously assessed and controlled; such studies may enable us to relate EEG abnormalities to the neuropathology of AD. Another interesting line of research is to investigate the effect of medication and therapy on the EEG of AD patients (see, e.g., [60–65]).

References

1. Mattson, M.: Pathways towards and away from AD. Nature **430** (2004).
2. Shimokawa, A., et al.: Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. Brain Cogn. **47**(3) (2001) 423–433.
3. Weiner, W.M.: Editorial: imaging and biomarkers will be used for detection and monitoring progression of early Alzheimer's disease. J. Nutr. Health Aging **4** (2009) 332.
4. Jeong J.: EEG Dynamics in patients with Alzheimer's disease. Clin. Neurophysiol. **115** (2004) 1490–1505.
5. Baker, et al.: EEG patterns in mild cognitive impairment (MCI) patients. Open Neuroimag. J. 2 (2008) 52–55.
6. Besthorn, et al.: Discrimination of Alzheimer's disease and normal aging by EEG data. Electroencephalogr. Clin. Neurophysiol. **103** (1997) 241–248.
7. van der Hiele, et al.: EEG correlates in the spectrum of cognitive decline. Clin Neurophysiol. **118**(9) (2007) 1931–1939.
8. Czigler, et al.: Quantitative EEG in early Alzheimer's disease patients—power spectrum and complexity features. Int. J. Psychophysiol. **68**(1) (2008) 75–80.
9. Gianotti, et al.: Correlation between disease severity and brain electric LORETA tomography in Alzheimer's disease. Clin. Neurophysiol. **118**(1) (2007) 186–196.
10. Herrmann, et al.: Human EEG gamma oscillations in neuropsychiatric disorders. Clin. Neurophysiol. **116**(12) (2005) 2719–2733.
11. Latchoumane, et al.: Multiway analysis of Alzheimers disease: Classification based on space-frequency characteristics of EEG time series, 2008. Proceedings of the World Congress on Engineering (2008).

12. Vialatte, et al.: Why sparse bump models? OHBM meeting, Melbourne, Australia, June 15–19 2008. Neuroimage 41(S1) (2008) S159.
13. Vialatte, et al.: Early diagnosis of Alzheimer's disease by blind source separation, time frequency representation, and bump modeling of EEG signals. Proceedings of the 15th International Conference on Artificial Neural Networks, Springer, LNCS, Vol. 3696 (2005) 683–692.
14. Schreiter, et al.: Alzheimer disease versus mixed dementias: an EEG perspective. Clin. Neurophysiol. **119**(10) (2008) 2255–2259.
15. Hornero, et al.: Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. Philos. Trans. R. Soc. A 367 (2009) 317–336.
16. Besthorn, et al.: Parameters of EEG dimensional complexity in Alzheimer's disease. Electroencephalogr. Clin. Neurophysiol. 95 (1995) 84–89.
17. Woon, et al.: Techniques for early detection of Alzheimer's disease using spontaneous EEG recordings. Physiol. Meas. 28 (2007) 335–347.
18. Babiloni, et al.: White-matter lesions along the cholinergic tracts are related to cortical sources of EEG rhythms in amnesic mild cognitive impairment. Hum. Brain Mapp. **30**(5) (2009) 1431–1443.
19. Babiloni, et al.: Directionality of EEG synchronization in Alzheimer's disease subjects. Neurobiol. Aging 30(1) (2009) 93–102.
20. Kramer, et al.: Synchronization measures of the scalp EEG can discriminate healthy from Alzheimer's subjects. Int. J. Neural Syst. 17 (2007) 1–9.
21. Dauwels J., Vialatte F., Rutkowski T., Cichocki A.: Measuring neural synchrony by message passing. Adv. Neural Inf. Process. Syst. 20 (NIPS 20), (2007) In press.
22. Dauwels, et al.: Quantifying statistical interdependence by message passing on graphs, Part I: One-dimensional point processes. Neural Comput. **21** (2009) 8, In press.
23. Dauwels, et al.: Quantifying statistical interdependence by message passing on graphs, Part II: Multi-dimensional point processes. Neural Comput. 21 (2009) 8, In press.
24. Dauwels, et al.: Quantifying the similarity of multiple multi-dimensional point processes by integer programming with application to early diagnosis of Alzheimers disease from EEG. Proceedings of the 30th IEEE EMBC (2008)
25. Dauwels, et al.: A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. NeuroImage (2009) In press.
26. Stevens, et al.: Dynamic regulation of EEG power and coherence is lost early and globally in probable DAT. Eur. Arch. Psychiatry Clin. Neurosci. **251** (2001) 199–204.
27. Koenig, et al.: Decreased EEG synchronization in Alzheimer's disease and mild cognitive impairment. Neurobiol. Aging. **26** (2005) 165–171.
28. Stam, et al.: Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. **28**(11) (2007) 1178–1193.
29. Pijnenburg, et al.: EEG Synchronization likelihood in MCI and AD during a working memory task. Clin. Neurophysiol. **115**(2004) 1332–1339.
30. Wan, et al.: Linear and nonlinear quantitative EEG analysis. IEEE Eng. Med. Biol. Mag. **27**(5) (2008) 58–63.
31. Park, et al.: Decreased EEG synchronization and its correlation with symptom severity in Alzheimer's disease. Neurosci Res. **62**(2) (2008) 112–117.
32. Jeong, J., Gore, J., Peterson, B.: Mutual information analysis of the EEG in patients with Alzheimer's disease. Clin. Neurophysiol. **112** (2001) 827–835.
33. Knyazeva, et al.: Topography of EEG multivariate phase synchronization in early Alzheimer's disease. Neurobiol. Aging. (2008)
34. Stam, et al.: Small-world networks and functional connectivity in Alzheimer's disease. Cereb. Cortex **17**(1) (2007) 92–99.
35. He Y., et al.: Neuronal networks in Alzheimer's disease. Neuroscientist (2009)

36. Jiang, Z.Y.: Abnormal corticol functional connections in AD: Analysis of inter- and intra-hemispheric EEG coherence. J. Zhejiang Univ. Sci. B **6** (2005) 259–264.
37. Storey J.: A direct approach to false discovery rates. J R Stat Soc Series B Stat Methodol **64**(3) (2002) 479–498.
38. Goldberger, et al.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. (2000)
39. Sotero, R.C., Trujillo-Barreto, N.J.: Biophysical model for integrating neuronal activity, EEG, fMRI and metabolism. Neuroimage **39**(1) (2008) 290–309.
40. Babiloni, et al.: Hippocampal volume and cortical sources of EEG alpha rhythms in mild cognitive impairment and AD. Neuroimage **44**(1) (2009) 123–135.
41. Moretti, et al.: Increase of theta/gamma and alpha3/alpha2 ratio is associated with amygdalo-hippocampal complex atrophy. J Alzheimers Dis. (2009)
42. Teipel, et al.: Regional networks underlying interhemispheric connectivity: An EEG and DTI study in healthy ageing and amnestic mild cognitive impairment. Hum. Brain Mapp. **30**(7) (2008) 2098–2119.
43. Julkunen, et al.: Navigated TMS combined with EEG in MCI and AD: A pilot study. J. Neurosci. Methods **172**(2) (2008) 270–276.
44. Musha, et al.: A new EEG method for estimating cortical neuronal impairment that is sensitive to early stage Alzheimer's disease. Clin. Neurophysiol. **113**(7) (2002) 1052–1058.
45. Babiloni, et al.: Homocysteine and electroencephalographic rhythms in Alzheimer disease: A multicentric study. Neuroscience **145**(3) (2007) 942–954.
46. Liddell, et al.: Rates of decline distinguish Alzheimer's disease and mild cognitive impairment relative to normal aging: Integrating cognition and brain function. J. Integr. Neurosci. **6**(1) (2007) 141–174.
47. Moretti, et al.: Increase of theta/gamma ratio is associated with memory impairment. Clin. Neurophysiol. **120**(2) (2009) 295–303.
48. Giannakopoulos, et al.: Electrophysiological markers of rapid cognitive decline in mild cognitive impairment. Front Neurol. Neurosci. **24** (2009) 39–46.
49. Jaeseung J., et al.: Nonlinear dynamical analysis of the EEG in patients with Alzheimer's disease and vascular dementia. J. Clin. Neurophysiol. **18**(1) (2001) 58–67.
50. Andersson, et al.: Electroencephalogram variability in dementia with lewy bodies, Alzheimer's disease and controls. Dement. Geriatr. Cogn. Disord. **26**(3) (2008) 284–90.
51. Gawel, et al.: The value of quantitative EEG in differential diagnosis of Alzheimer's disease and subcortical vascular dementia. J. Neurol. Sci. (2009)
52. Bonanni, et al.: EEG comparisons in early Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up. Brain **131** 690–705.
53. Uhlhaas, et al.: Visual perceptual organization deficits in Alzheimer's dementia. Dement. Geriatr. Cogn. Disord. **25**(5) (2008) 465–475.
54. Papaliagkas, et al.: Usefulness of event-related potentials in the assessment of mild cognitive impairment. BMC Neurosci. **9** (2008) 107.
55. van Deursen, et al.: 40-Hz steady state response in Alzheimer's disease and mild cognitive impairment. Neurobiol Aging (2009).
56. Caravaglios, et al.: Decreased amplitude of auditory event-related delta responses in Alzheimer's disease. Int. J. Psychophysiol. **70**(1) (Oct 2008) 23–32.
57. van Deursen, et al.: Response speed, contingent negative variation and P300 in Alzheimer's disease and MCI. Brain Cogn. **69**(3) (2009) 592–599.
58. Güntekin, et al.: Decrease of evoked delta, theta and alpha coherences in Alzheimer patients during a visual oddball paradigm. Brain Res. **1235** (2009) 109–116.
59. Bobkova, et al.: Interhemispheric EEG differences in olfactory bulbectomized rats with different cognitive abilities and brain beta-amyloid levels. Brain Res. (2008).
60. Babiloni, et al.: Ibuprofen treatment modifies cortical sources of EEG rhythms in mild Alzheimer's disease. Clin. Neurophysiol. **120**(4) (2009) 709–718.

61. Gianotti, et al.: Rivastigmine effects on EEG spectra and three-dimensional LORETA functional imaging in Alzheimer's disease. Psychopharmacology (Berl) **198**(3) (2008) 323–332.
62. Yener, et al.: Increased frontal phase-locking of event-related theta oscillations in Alzheimer patients treated with cholinesterase inhibitors. Int. J. Psychophysiol. **64**(1) (2007) 46–52.
63. Kogan, et al.: Long-term tetrahydroaminoacridine treatment and quantitative EEG in Alzheimer's disease. J. Neural Transm. Suppl. (72) (2007) 203–6.
64. Babiloni, et al.: Donepezil effects on sources of cortical rhythms in mild Alzheimer's disease: Responders vs. non-responders. Neuroimage **31**(4) (2006) 1650–1665.
65. Moraes, et al.: The effect of donepezil on sleep and REM sleep EEG in patients with Alzheimer disease: a double-blind placebo-controlled study. Sleep **29**(2) (2006) 199–205.
66. Park, et al.: Multiscale entropy analysis of EEG from patients under different pathological conditions. Fractals **15**(4) (2007) 399–404.

Analysis of EEG Time Series Recorded from Alzheimer Patients Based on Their Spectral Content

Aurélien Hazart, François-Benoît Vialatte, and Andrzej Cichocki

Abstract In this paper, EEG recordings on each channel are seen as non stationary time series. We assume a piecewise stationary model for the signal where the changes affect the power spectrum. Such a model is particularly relevant for analyzing Alzheimer disease (AD) dataset as the disease leads to variation of the frequency rhythms. Our method is devoted to the estimation of the partition. We define a criterion based on a regularized periodogram to estimate the spectrum and on the power spectrum on predefined frequency bands to estimate the change points. The method produces new markers designed to be used for the discrimination between control and AD patients.

1 Introduction

Studying Alzheimer disease with Electroencephalogram (EEG) recordings has encountered a lot of interest in the past decades. One advantage of EEG imaging technique is the high resolution in time, that allows also a high resolution in frequency. Numerous authors take advantage of this characteristic to analyze Alzheimer dataset. In [1], the author describes the relation between abnormal EEG dynamics and AD. Briefly, it is generally admitted that AD patients show a slowing down of the brain activity (shift to lower frequencies) and a decrease of coherence in the higher frequencies.

An important issue is to classify patients in different groups, namely AD (Alzheimer Disease), MCI (Mild cognitive impairment) and CTR (age-matched control). The goal is to achieve detection of Alzheimer disease in early stage (MCI). Several works propose to discriminate AD, MCI and CTR patients with frequency-based approaches (see the review [1]). Most of them are based on representation of

A. Hazart (✉)
Laboratory for Advanced Brain Signal Processing, Riken Brain Science Institute, Wakoshi, Saitama, Japan
e-mail: auhazart@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_107,

the signal in the frequency domain [2], in the frequency-spatial domain [3] or in the time-frequency domain [4]. Even if they obtain good classification ratio of AD and MCI patients and sometimes of MCI and CTR patients, the problem is still open.

In this paper, we base our method on a time-frequency analysis but without an explicit time-frequency representation. We aim at partitioning the time series recorded on each channel in segments with similar spectral characteristics. We estimate the time points of the ruptures between stationary segments. Time segmentation of EEG signal is already applied in the literature but only for partitions based on spatially stationary map. For instance in [5], the Global Dissimilarity measures the stability of the spatial maps over time (microstate) and is applied for AD dataset in [6]. In our approach, we prefer to follow the idea of [4] that does not use the spatial information. Each channel is partitioned separately according to the power spectrum in specific frequency bands.

2 Method

2.1 Model

Several methods exist in the literature for the change point detection problem. Our method belongs to the off-line parametric multiple change point class: All the signal is used to estimate multiple changes that affect latent parameters.

We note $y_n, n = 1, \ldots, N$ the signal recorded by one channel. We assume that $K - 1$ change points divide the signal into K segments. The sorted change times are noted $t_k, k = 1, \ldots, K - 1$ with the extension $t_0 = 1$ and $t_K = N$. The signal on segment k corresponds to the vector $[y_{t_{k-1}+1}, \ldots, y_{t_k}]$ and the elements are noted y_{m_k}, $m_k \in]t_{k-1}, t_k]$. To model the spectral distribution of each segment, we simply compute the discrete Fourier transform (DFT) for P discrete frequencies as follow:

$$y_{m_k} = \frac{1}{P} \sum_{p=0}^{P-1} a_{p,k} \exp\{2j\pi m_k p/P\} + e_{m_k} \tag{1}$$

where the $a_{p,k}$ denote the unknown DFT of y_{m_k} and $\mathbf{e} = [e_1, \ldots, e_n]$ is the noise term, assumed to be a vector of i.i.d. zero mean Gaussian samples. Eq. (1) can be rewrite in vector form as $\mathbf{y}_k = \mathbf{H}\mathbf{a}_k + \mathbf{e}_k$. The vectors $\mathbf{y}_k$ and $\mathbf{e}_k$ contain the data and the noise of segment k respectively, $\mathbf{H}$ the exponential terms $1/P \exp\{2j\pi m_k p/P\}$ and $\mathbf{a}_k$ the DFT. Assuming independence of the observations between segments, the estimation of the $\mathbf{a}_k$ under the maximum likelihood principle leads to the following criterion (to be minimized in $\mathbf{a}_k$ and t_k):

$$J(t_k, \mathbf{a}_k) = \sum_{k=1}^{K} J_k(\mathbf{a}_k) \quad \text{with} \quad J_k(t_k, \mathbf{a}_k) = (\mathbf{y}_k - \mathbf{H}\mathbf{a}_k)^{\dagger}(\mathbf{y}_k - \mathbf{H}\mathbf{a}_k). \tag{2}$$

To find the change points t_k and eventually the coefficients $\mathbf{a}_k$ that ensure homogeneous spectrum within each segment, we consider constraints on the $\mathbf{a}_k$.

First, we add a separable quadratic regularization term $\mathbf{a}_k^\dagger \mathbf{a}_k$ [7] to avoid irregular spectrum. Then, following several works on EEG [3, 8], we suppose the changes affect power spectrum in specific frequency bands, designed by the index $i = 1, \ldots, I$. To maximize the power spectrum computed per bands and per segment, we penalize the opposite of the power spectrum in each bands:

$$J_k(t_k, \mathbf{a}_k) = (\mathbf{y}_k - \mathbf{H}\mathbf{a}_k)^\dagger (\mathbf{y}_k - \mathbf{H}\mathbf{a}_k) + \lambda_1 \mathbf{a}_k^\dagger \mathbf{a}_k - \lambda_2 \sum_{i=1}^{I} \left(\sum_{p_i} |a_{k,\,p_i}|^2 \right)^2 \tag{3}$$

where λ_1, λ_2 are positive regularization parameters, $.^\dagger$ is the hermitian operator ($\mathbf{a}_k$ is a vector of complex) and the p_i denotes the discrete frequencies of band i.

With the previous criterion (2)–(3), we assume that the number of segments K is known. The estimation of K requires a model selection procedure. The standard method is to minimize the criterion $J(K) = J(\hat{n}_k, \hat{\mathbf{a}}_k) + \lambda_3 \phi(K)$ where $\phi(K)$ penalizes the number of segments to avoid over segmentation, for instance $\phi(K) = 2K$ (Bayesian Information Criterion, BIC).

2.2 Computation

Ideally, the minimization of the criterion (2) and (3) in $t_k, \mathbf{a}_k$ would give the estimated time change points as well as the estimated spectral content per segment. However regarding the difficulty of the joint estimation problem, we decompose the problem in two steps. We first estimate the coefficients $\hat{\mathbf{a}}_k$ by minimizing the two first terms of (3), leading to the explicit form [7]:

$$\hat{\mathbf{a}}_k = (1/P + \lambda_1)^{-1} \mathbf{H}^\dagger \mathbf{y}_k \tag{4}$$

which corresponds to the periodogram regularized with a quadratic separable term. Advantage of the regularized periodogram is a better spectral estimation when the size of a segment is small. Then we minimize the third term of (3) replacing $a_{k,i}$ by their previous estimation: $-\lambda_2 \sum_i (\sum_p |\hat{a}_{k,i}|^2)^2$. As in [8], the optimization is achieved with a dynamic programming algorithm [9]. With this approximation, our method can be seen as a continuation of [8] in the following ways: The spectrum estimation is done with a regularized periodogram and we apply the method on a real EEG AD database.

3 Results

The method is applied on real EEG recordings. The database was recorded in Japan from 83 patients in a eyes-closed rest state. Even without specific stimuli, the spontaneous activity of the brain shows a succession of pattern activations, and the analyze on the signal in the course of time remains relevant. The database consists of 20 s

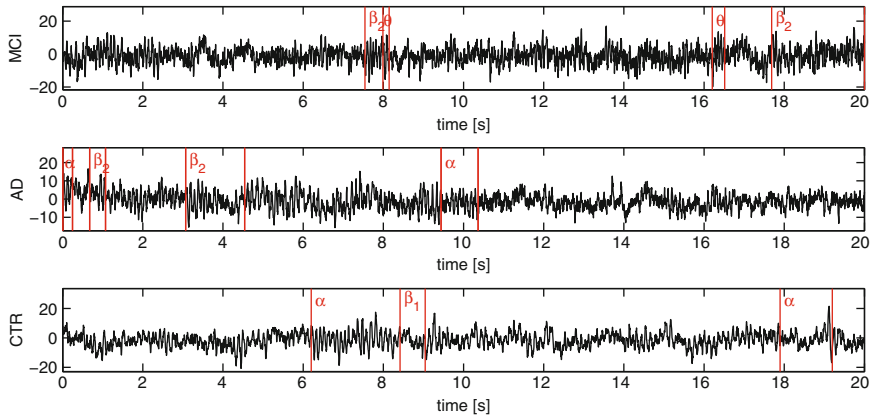


Fig. 1 Segmentation and corresponding dominant frequency of the EEG recordings at channel P4 from three patients (MCI, AD, CTR)

continuous artifact-free observations recorded on 21 channels at 200 Hz from three groups of patients (AD, MCI, CTR). Following [3] which show significant changes in theta, alpha and beta2 frequency bands, in centro-parietal area, we focus here on these frequency bands (δ: 4–8 Hz, α: 8–12 Hz, β_1: 12–15 Hz, β_2: 15–19 Hz) and location.

Qualitatively, the results obtained on all the database tend to show that segments with a dominant frequency are more present in MCI patients than in AD or CTR. Also, the length of the segments is longer for healthy patients than for AD or MCI, and is usually shorter in θ band (less than 0.2 s) than in β or α (few seconds). We find that the dominant frequency band for MCI patients are more often θ and β_2, which corroborate previous works on the early stage of AD [1, 3]. Because of space limitation, we present the segmentation result on one channel, the P4 electrodes (right parietal location, 10–20 system) and for three patients randomly selected (Fig. 1). For each patient, the algorithm found three or four segments showing a dominant frequency, with lengths comprising between 200 ms and 2 s. Based on these results, a potential relevant feature for AD detection is the comparison of the power spectrum per frequency band weighted by the length of the corresponding segments.

4 Conclusion

We proposed here a method for partitioning EEG time series according to their spectral distribution. It is an extension of classic spectral method, as it also considers time location of frequency. We show that the method is a good candidate for analyzing AD database. Next steps would be to merge the segmentation of multiple channels and to extract relevant markers for AD classification like temporal and spatial position of frequency or number and length of segments.

References

1. Jeong, J.: Eeg dynamics in patients with alzheimer's disease. Clini. Neurophysiol. **115**(7) (2004) 1490–1505.
2. Coben, L., Danziger, W., Berg, L.: Frequency analysis of the resting awake eeg in mild senile dementia of alzheimer type. Electroencephalogr. Clin. Neurophysiol. **55**(4) (Apr 1983) 372–380
3. Huang, C., Wahlund, L.O., Dierks, T., Julin, P., Winblad, B., Jelic, V.: Discrimination of alzheimer's disease and mild cognitive impairment by equivalent EEG sources: A cross-sectional and longitudinal study. Clin. Neurophysiol. **111** (2000) 1961–1967
4. Vialatte, F., Cichocki, A., Dreyfus, G., Rutkowski, T.M.T., Gervais, R.: Blind source separation and sparse bump modelling of time frequency representation of EEG signals: New tools for early detection of alzheimer's disease. IEEE Workshop on Machine Learning for Signal Processing 2005 (MLSP'05). (Sep 2005) 27–32
5. Lehmann, D., Skrandies, W.: Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. **6**(48) (Jun 1980) 609–621
6. Ihla, R., Dierksb, T., Froelichb, L., Martinb, E.M., Maurerb, K.: Segmentation of the spontaneous eeg in dementia of the alzheimer type. Neuropsychobiology **27** (1993) 231–236
7. Sacchi, M., Ulrych, T., Walker, C.: Interpolation and extrapolation using a high-resolution discrete fourier transform. IEEE Trans. Signal Process. **46**(1) (Jan 1998) 31–38
8. Lavielle, M.: Using penalized contrasts for the change-point problem. Signal Process. **85**(8) (2005) 1501–1510
9. Auger, I., Lawrence, C.: Algorithms for the optimal identification of segments neighborhoods. Bull. Math. Biol. **51** (1989) 39–54

Modeling Transient Oscillations in the EEG of Patients with Mild Cognitive Impairment

François-Benoît Vialatte, Jordi Solé-Casals, Aurélien Hazart, David Prvulovic, Justin Dauwels, Johannes Pantel, Corinna Haenschel, and Andrzej Cichocki

Abstract We explore the potential of transient local synchrony in EEG, as a marker for MCI (mild cognitive impairment). EEG signals of patients with MCI are transformed to a wavelet time-frequency representation, and afterwards a sparsification process (bump modeling) extracts time-frequency oscillatory bursts. We observed that organized oscillatory events contain stronger discriminative signatures than averaged spectral EEG statistics for patients in a probable early stage of Alzheimer's disease. Specifically, bump modeling enhanced the difference between MCI patients and age-matched control subjects in the θ and high β frequency ranges. This effect is consistent with previous results obtained on other databases.

1 Introduction

Alzheimer's disease is a brain disorder, whose prevalence is dramatically increasing (due to the general increase of life expectancy), threatening our societies. It would be a great asset if we were able to detect it as early as possible, in particular, before the onset of AD, a stage called predementia or mild cognitive impairment (MCI). However, MCI has few, unreliable symptoms. A cost-efficient technique would be necessary to screen populations at risk, potentially thousands to even millions of people. Electroencephalography (EEG) is cost-effective, and was suggested as a tool for diagnosing AD [1]; however its specificity to the disease is low, so that its reliability is sometimes questioned. Nevertheless, EEG data are not totally exploited by medical teams, especially its main advantage: a very precise time resolution, allowing investigations of brain dynamics (see [1] for a review). Our purpose here is to explore the potential of transient local synchrony in EEG, as a marker for the MCI stage.

F.-B. Vialatte (✉)
Laboratory for Advanced Brain Signal Processing, RIKEN BSI, Saitama, Japan
e-mail: fvialatte@brain.riken.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_108,

Brain signals evolve quickly and non-linearly in time. EEG recordings consist of stable and sustained activities on one hand, and unstable and transitory activities on the other hand. When these transitory activities become organized in patterns, they form bursts in EEG signals. Here, we are interested in these bursts, usually characterized as a succession of 4–5 oscillations, within a limited frequency range: they are hypothesized to be representative of transient synchronies of neural populations [2, 3]. In the past few years, a lot of attention was devoted to signals evoked or induced by specific stimulations. The brain responses to stimuli can be observed in EEG signals, and such oscillatory bursts can be part of this response. But this kind of activity is not shut down during rest periods. We intend here to study oscillatory bursts in EEG signals recorded in rest condition, at the single trial level (not using averaging). For this purpose, we applied bump modeling.

2 Methods

The database was recorded from 22 patients with MCI and 31 age-matched controls. A 64 channel electrode-cap was fitted to the participants's head with the ground electrode at position AFz and the reference electrode at position FCz. An additional vertical electrooculogram electrode was placed below the right eye. Electrode impedance was kept below 5 kΩ. Recording, digitization and preprocessing of the EEG data were carried out with a BrainAmp amplifier and the BrainVision Recorder software (Brain Products, Munich, Germany). Recording was done in the Laboratory for Neurophysiology and Neuroimaging, Department of Psychiatry, Johann Wolfgang Goethe University, 60590 Frankfurt, Germany in collaboration with the Max Planck Institute for Brain Research, 60528 Frankfurt, Germany.All participants provided informed consent prior to the study. Ethical approval was obtained from the ethics committee of the Medical School, Johann Wolfgang Goethe University, Frankfurt am Main, Germany.

Sampling rate was 500 Hz, and offline digital filters (Butterworth, zero-phase forward and reverse) were set with bandpass at 0.5–90 Hz, and notch at 50 Hz. ICA pre-processing of data was performed using ICAlab ver 3.0, with the IWASOBI algorithm. Time-frequency sparse modeling was performed using Matlab ® 7.0, and the ButIf toolbox [3]. Three EEG researchers visually inspected EEGs, and each recording's least corrupted continuous 20 s interval (from 1 min recorded in rest with eyes closed condition) were chosen for the analysis. Each trial was then decomposed using ICA , and artifacts (drifts, eye blinks, sharp waves, abnormal amplitude sources) were removed independently using the procedure detailed in [4]. The bump modeling method is covered in detail in [2, 3]. EEG signals were first transformed to a time-frequency representation using complex Morlet wavelets. The wavelet map was computed from 4 to 85 Hz with 1 Hz steps, and afterwards normalized using a z-score normalization (see e.g. [5]). We then modeled the oscillatory patterns present in the normalized time-frequency map:

$$E(A,h,w,f,t,y,x) = \sum_{x=1}^{W} \sum_{y=1}^{H} \left\| \omega_{y,x}(s,\tau) - \xi(A,h,w,f,t,y,x) \right\|^2 . \qquad (1)$$

Where the sums are computed over a set of windows, $\xi(A, h, w, f, t, y, x)$ are half-ellipsoid functions, y and x are respectively the time and frequency position of the adaptation window on the time-frequency map (fixed parameters), f and t are respectively the time and frequency position of the bump on the time-frequency map, h and w are respectively the height and width of the bump, A is the amplitude of the bump, and $\|.\|$ is a Frobenius norm. Adaptation is performed using a combination of first and second order gradient descent. The results presented here were obtained with a pruning threshold $F_t = 0.30$, and a z-score offset $\phi = 0$.

3 Results and Discussion

We compared the statistics of EEG power before (using the wavelet map before and after z-score) and after bump modeling (using the amplitude A of the bumps in the frequency range), in six frequency ranges: θ (4–8 Hz), α_1 (8–10 Hz), α_2 (10–12 Hz), β_1 (12–19 Hz), β_2 (19–25 Hz), and γ (25–45 Hz). These results are computed in relative power (the power in each frequency range is divided by the total power). First, we compared the general average over all electrodes, for each of the six frequency ranges, using a Mann-Whitney test[1] (Table 1). The difference between the two groups is increased after modeling, and significant effects can be observed in the θ and β_2 ranges, even after statistical correction.

Second, we located the effect by averaging the amplitude of bumps in nine scalp zones (in Table 2), as was done in our previous publications (e.g. [6]). Because we already know that θ and β_2 ranges have significant effects (see Table 1), the effects in these ranges can be interpreted without post-hoc correction (interpreted as a localization of the significant effect). The difference in the θ range appears to be independent on location, whereas the effect in the β_2 only appears in posterior locations (central, parietal right and left, and occipital). The differences in the other frequency ranges were analyzed after post-hoc correction. Significant differences appear in the α_1 range in prefrontal channels, and in the α_2 range in several areas,

Table 1 Difference between MCI patients and Control subjects, before and after the wavelet map is modeled with bumps (relative power)

Frequency range (Hz)	Wavelets p-value	Bumps p-value
θ (4–8)	0.2014	**0.013**[a]
α_1 (8–10)	0.439	0.661
α_2 (10–12)	0.634	0.054
β_1 (12–19)	0.309	0.376
β_2 (19–25)	0.702	**0.0035**[a]
γ (25–45)	0.356	0.634

[a] Significant effects. Significant effects after post-hoc correction are shown in bold.

[1] Wavelet coefficients are usually not distributed according to a normal distribution, hence a non-parametric test has to be used.

Table 2 Difference between MCI patients and Control subjects, after the wavelet map is modeled with bumps (relative power). The difference between the two groups is shown depending on 9 scalp zones

Range (Hz)→ Area ↓	θ (4–8)	α_1 8–10)	α_2 (10–12)	β_1 (12–19)	β_2 (19–25)	γ (25–45)
Prefrontal	$\mathbf{1.1 \cdot 10^{-8}}^a$	$\mathbf{1.4 \cdot 10^{-3}}^a$	$\mathbf{9.8 \cdot 10^{-7}}^a$	0.23	0.12	0.83
Left frontal	$\mathbf{3.7 \cdot 10^{-7}}^a$	0.01[a]	0.0010[a]	0.77	0.22	0.83
Right frontal	$\mathbf{1.2 \cdot 10^{-8}}^a$	0.023[a]	0.0063[a]	0.57	0.032	0.48
Left temporal	$\mathbf{2.0 \cdot 10^{-3}}^a$	0.088[a]	0.018[a]	0.42	0.70	0.58
Central	$\mathbf{1.0 \cdot 10^{-5}}^a$	0.54	0.43	0.13	0.096	0.43
Right temporal	$\mathbf{8.7 \cdot 10^{-9}}^a$	0.041[a]	$\mathbf{2.1 \cdot 10^{-3}}^a$	0.87	0.039[a]	0.44
Left parietal	$\mathbf{1.0 \cdot 10^{-5}}^a$	0.010[a]	0.23	0.25	0.26	0.93
Right parietal	$\mathbf{9.1 \cdot 10^{-8}}^a$	0.013[a]	$\mathbf{2.6 \cdot 10^{-3}}^a$	0.76	0.066	0.041[a]
Occipital	$\mathbf{6.7 \cdot 10^{-7}}^a$	0.31	$\mathbf{5.1 \cdot 10^{-6}}^a$	0.16	0.054	0.73

[a] Significant effects after post-hoc correction are shown in bold.

especially over the right hemisphere. As a comparison, power computed directly from wavelet coefficients did not show any significant effect, even without post-hoc correction (table not shown due to space limitation).

Classification rate in the θ range was compared, using either the amplitude on the wavelet map, or from the bump modeling. We used a linear classifier (linear discriminant analysis) in a leave-one-out scheme. The θ power, computed for each patient, was used as an input feature. The classification error decreased from 50.0% (wavelets) to 34.6% (bumps). Similarly, using the β_2 range, the classification error decreased from 50.0% (wavelets) to 36.5% (bumps). Classification using other frequency ranges remained stable (i.e. inaccurate) after bump modeling. A better classification rate should be obtained with optimized parameter combinations (using feature selection) together with a more complex classifier; this is however out of the scope of the present investigation.

We observed that organized oscillatory events contain stronger discriminative signatures of MCI than averaged spectral EEG statistics. Similarly to our previous results on patients with MCI, using another database of patients in MCI stage *vs.* Control subjects [6], bump modeling improved the separation between MCI patients and control subjects, specifically in the θ and β_2 ranges. We report that the θ range activity of transient oscillations increased significantly in temporal areas for patients with MCI, while β_2 range activity decreased. We were thus able to confirm here our previous observation, using a different database.

Background activity in EEG is mostly attributed to cortical neural events; on the other hand, the oscillatory bursts, generated by locally synchronous neural populations, could be related to inter-area interactions, including sub-cortical areas. Indeed, low-frequency synchrony is probably representative of cortico-subcortical connectivity [7]. Subcortical damages are induced in the early stage of Alzheimer's disease, and have been correlated with low-frequency power changes [8, 9]. As a conclusion, using bump modeling allowed us to classify the two groups specifically in the

θ and β range. We postulate that the observed strong increase of θ range transient oscillatory activity could be a correlate of sub-cortical damages.

References

1. Jeong, J.: EEG dynamics in patients with alzheimer's disease. Clin. Neurophysiol. **115** (2004) 1490–1505.
2. Vialatte, F., Martin, C., Dubois, R., Haddad, J., Quenet, B., Gervais, R., G, D.: A machine learning approach to the analysis of time-frequency maps, and its application to neural dynamics. Neural Netw. **20**(2007) 194–209.
3. Vialatte, F., Solé-Casals, J., Dauwels, J., Maurice, M., Cichocki, A.: Bump time-frequency toolbox: a toolbox for time-frequency oscillatory bursts extraction in electrophysiological signals. BMC Neurosci. **10**(46) (2009).
4. Vialatte, F.B., Solé-Casals, J., Maurice, M., Latchoumane, C., Hudson, N., Wimalaratna, S., Jeong, J., Andrzej, C.: Improving the quality of EEG data in patients with alzheimers disease using ica. In: Proceedings of the 15th International Conference on Neural Information Processing, ICONIP, Auckland, New Zealand, November 25–28 2008 – LNCS, Part II, 5507 (2009). 979–986.
5. Vialatte, F., Solé-Casals, J., Cichocki, A.: EEG windowed statistical wavelet scoring for evaluation and discrimination of muscular artifacts. Physiol. Meas. **29**(12) (2008) 1435–1452.
6. Vialatte, F., Maurice, M., Cichocki, A.: Why sparse bump models? In: Proceedings of OHBM Meeting: June 15–19 2008, Melbourne, Australia – Neuroimage, Vol. 41(S1) (2008) S159.
7. Uhlhaas, P., Singer, W.: Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. Neuron **52** (2006) 155–168.
8. Helkala, E., Hänninen, T., Hallikainen, M., Könönen, M., Laakso, M., Hartikainen, P., Soininen, H., Partanen, J., Partanen, K., Vainio, P., Riekkinen, P.S.: Slow-wave activity in the spectral analysis of the electroencephalogram and volumes of hippocampus in subgroups of alzheimer's disease patients. Behav. Neurosci. **110**(6) (1996) 1235–1243.
9. Fernández, A., Arrazola, J., Maestü, F., Amo, C., Gil-Gregorio, P., Wienbruch, C., Ortiz, T.: Correlations of hippocampal atrophy and focal low-frequency magnetic activity in alzheimer disease: Volumetric mr imaging-magnetoencephalographic study. AJNR Am. J. Neuroradiol. **24**(3) (2003) 481–487.

Discrimination of the Sleeping EEG by the Adaptive Rhythmic Component Extraction

Yohei Tomita, Hironobu Fukai, Yasue Mitsukura, Toshihisa Tanaka, and Jianting Cao

Abstract It is crucial to practically estimate the driver dozing. For this purpose, the analysis of the frequency components of the electroencephalogram (EEG) is accepted as the most favorable. Therefore, we propose to extract the frequency components of the EEG, by using the rhythmic component extraction (RCE), proposed by Tanaka et al. First of all, we analyze differences of power spectrum among the channels. As this result, the recognition accuracy between the sleeping and the waking when the frontal lobe is used is 76.9%. In addition, we distinct the EEG data by using weight features of the RCE. As results, the accuracy is up to 94.1%. It is shown the effectiveness of the weight analysis. Therefore, there is possibility that we can extract effective electrodes by the RCE weights.

1 Introduction

Driver dozing caused by the lack of the sleeping is partly due to the traffic accidents, and it has been more serious. The National Highway Traffic Safety Administration (NHTSA) conservatively estimates that 100,000 police-reported crashes are the direct results of driver fatigue each year. These crashes result in an estimated 1,550 deaths, 71,000 injuries, and $ 12.5 billion in monetary losses [1]. Therefore, the accurate and nonintrusive real-time monitoring of driver dozing would be highly desirable.

In order to judge whether the subject sleeps or not, the electroencephalogram (EEG) is reported to be suitable [2–4]. In computer simulations, many researchers analyze the frequency components of the EEG obtained by the frequency analysis (e.g. Fourier transform). The effectiveness of the frequency components were confirmed in our previous study [5]. Furthermore, Tanaka et al. pointed out if the frequency of interest is known in advance, it is more natural to directly estimate

Y. Tomita (✉)
Tokyo University of Agriculture and Technology, Koganei, Tokyo, 184-8588, Japan
e-mail: y_tomita@mi.lab.tuat.ac.jp

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_109,

a component with respect to the frequency range [6]. To estimate this component, the so-called rhythmic component extraction (RCE) method has been proposed. The extraction is made by combining multi-channel signals with weights that are optimally sought for such that the extracted component maximally contains the power in the frequency range of interest and suppresses that in unnecessary frequencies. Therefore, the frequency of interest which cannot be seen in each single channel is detectable by the RCE signal. In this study, we calculate the discrimination accuracy of the sleeping and before sleeping by each channel and the RCE signal. Furthermore, we analyze the RCE performance by changing the time window (the rectangular, the Hann, and the Blackman).

2 Rhythmic Component Extraction

RCE is a method to extract a certain frequency range of the EEGs. It is expressed as

$$\hat{x}_{(n)}[k] = \sum_{i=0}^{M} w_{i,(n)} x_{i,(n)}[k]. \tag{1}$$

Observed signal $x_{i,(n)}[k](k = 0, ..., N-1)$ is based on the channel $i(i = 1, ..., M)$ in frame n. $w_{i,(n)}$ is the weight determined to maximize the power of specific frequency component, whereas the power of the other frequency component is minimized. $\Omega_1 \subset [0, \pi]$ is the frequency components we want to extract, and $\Omega_2 \subset [0, \pi]$ is the other frequency components we want to suppress, when the Fourier transform is computed for $x_{i,(n)}[k]$. Therefore, the cost function given as

$$J_{1,(n)}[w] = \frac{\int_{\Omega_1} |\hat{X}_{(n)}(e^{-j\omega})|^2 d\omega + \epsilon |r_{(n)}|^2}{\int_{\Omega_2} |\hat{X}_{(n)}(e^{-j\omega})|^2 d\omega}. \tag{2}$$

is maximized. Let $\hat{X}_{(n)}(e^{-j\omega})$ be the discrete-time Fourier transform (DTFT) of $\hat{x}_{(n)}[n]$. $r_{(n)}$ and ϵ serve as adaptation of the RCE. We take into consideration of the correlation between the signals extracted in the previous frame $n-1$ and that in the current frame n. That correlation $r_{(n)}$ is given as follows:

$$r_{(n)} = w_{(n)}^T X_{(n)} X_{(n-1)}^T w_{(n-1)}, \tag{3}$$

where $w_{(n)} = [w_{1,(n)}, ..., w_{M,(n)}]$, and $[X_{(n)}]_{ik} = x_{i,(n)}[k]$.

3 Experimental Procedure

We used the NEUROSCAN system for measurement. There are 32 channels (including 2 reference electrodes), and sampling frequency is 1,000 Hz. In the experiment, we investigate the difference of the EEG between before and just after the sleeping.

There is not clear definition of the dozing, so that we define just after the sleeping as following. After the attachment of the electroencephalograph, we measure the EEG for 1 min. We define the data as before sleeping. After that, the subject keeps relaxing and sleeps. For determination that the subject sleeps, we drum on the table softly by a finger every 4 min. If he notices, the subject moves his fingers in small motions, and it is defined that the subject is not sleeping. If he does not notice that sound, the subject cannot move fingers. When it happens the EEG is measured for 1 min. We define the data as during the sleeping.

4 Results and Discussions

In the conventional study, the theta and the alpha wave are known to change depending on the sleeping conditions. Therefore, we detect the theta and the alpha wave components by the RCE. To extract the theta wave component, Ω_1 is set to the range corresponding to 4–7 Hz, and Ω_1 is 8–13 Hz for the alpha wave component. The RCE signals are extracted from all 30 channels per 1 s. The sampling rate is down-sampled to 500 Hz.

First of all, for discrimination between before and during the sleeping, we pay attention to the power spectrum analyzed by the Fourier transform in the signal of each channel. The Fourier spectrum is calculated per 1 s for each signal. Moreover, we use distinction analysis using the support vector machine (SVM). Input vectors are the theta wave power and the alpha wave power. Discrimination by the SVM is carried out in leave-one-out cross-validation (LOOCV) per 1 s. As the result, the accuracy in central lobe is 64.2% by C3, CZ, C4, and the accuracy in the frontal lobe is 76.9% by FP1, FZ, FP2.

Moreover, we pay attention to the weight of each channel from the RCE signal. First of all, we define the weight of channel i at the time k as $w_{i,(n)}[k](theta)$ when we extract the theta wave component. On the other hand, we define $w_{i,(n)}[k](alpha)$ when we extract the alpha wave component. From these variables we define the weight feature as following.

$$W_{i,(n)}[k](alpha/theta) = w_{i,(n)}[k](alpha)/w_{i,(n)}[k](theta) \tag{4}$$

Next, we calculate the average of $W_{i,(n)}[k](alpha/theta)$ among 10 s, time shifting is 1 s. To see time variable of these features easily, we project these features to the 2D space by the principal component analysis (PCA). The variable is the weight of the channel. The result (rectangular window and $\epsilon = 0$) is shown in Fig. 1. These data could be separated. Moreover, we use the SVM for distinction analysis. Input data is $w_{i,(n)}[k](alpha/theta)$ among channels. Table 1 shows the accuracy of separation using the different window functions and ϵ. As the result, the high accuracy is confirmed by using the rectangular window. Therefore, we can advocate that the sleeping conditions are also can be detected by the RCE weight.

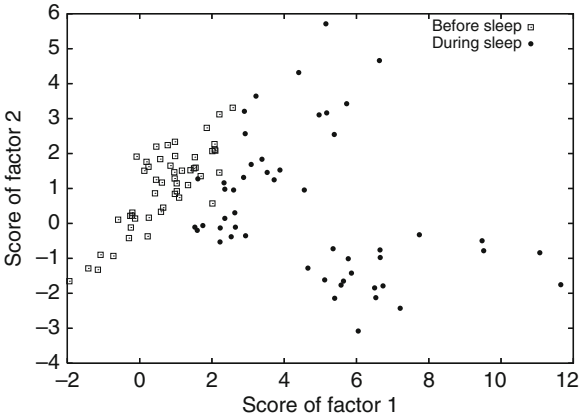


Fig. 1 Results of PCA mapping for the weights of all the channels

Table 1 The accuracy (%) of estimation of the sleeping conditions

ϵ / window function	Rectangular	Hann	Blackman
0	50.0	76.5	88.3
250	91.2	64.7	64.7
500	77.5	64.7	82.4
1,000	94.1	61.8	90.2

5 Conclusions

In order to analyze the sleeping EEG, we extract the theta and the alpha wave components by the RCE. Moreover, we used the SVM to discriminate the sleeping conditions. From this analysis, the high accuracy is confirmed using the channels in frontal lobe. Furthermore, to investigate the effectiveness of the RCE weights, we try to find the specific electrodes depending on the sleeping conditions. As the result, the accuracy is up to 94.1%. This shows the effectiveness of the RCE. In the future works, we analyze the important electrodes for the sleeping conditions estimation.

References

1. Royal, D.: Volume 1 Findings report; national survey on distracted and driving attitudes and behaviors, 2002. The Gallup Organization, Washington, DC, Technical Report DOT HS 809 566 (Mar 2003).
2. Golz, M., Sommer, D., Chen, M., Mandic, D., Trutschel, U.: Feature fusion for the detection of micro sleep events. J. VLSI Signal Process. **49** (2007) 329–342.
3. Gautama, T., Mandic, D.P., Van Hulle, M.M.: The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D **190** (2004) 167–176.

4. Gautama, T., Mandic, D.P., Van Hulle, M.M.: On the indications of nonlinear structures in brain electrical activity. Virtual J. Biol. Phys. Res. (Selected Papers from Phys. Rev. E) **5**(8) (2003).
5. Tanaka, T., Saito, Y.: Rhythmic component extraction for multi-channel EEG data analysis. Proceedings of the ICASSP 2008 (2008) 425–428.
6. Tomita, Y., Ito, S., Mitsukura, Y., Tanaka, T.: Rhythmic component extraction for sleeping electroencephalogram analysis. Proceedings of NCSP 2009 (2009) 443–446.

Analyzing EEG of Quasi-brain-death Based on Approximate Entropy Measures

Kun Yang, Qiwei Shi, Jianting Cao, Rubin Wang, Huili Zhu, and Zhaoshui He

Abstract Significant characteristic differences exist between the group of comatose patients and brain deaths. Statistical analysis methods have advantages in analyzing recorded EEG signals. In this paper, we apply a method based on approximate entropy (ApEn) associated with dynamic complexity to analyze a total of 35 patients. The experimental results illustrate effectiveness of the proposed method in measuring the EEG data and well performance in evaluating the differences between comatose patients and quasi-brain-deaths.

Keywords Electroencephalography (EEG) · Quasi-brain-death · Approximate entropy (ApEn) · Dynamic complexity

1 Introduction

The concept of brain death first appeared in 1960s. In clinical practice, brain death is defined as the absence and permanent loss of all brain and brain stem function [1]. Based on this definition, for example, the Japanese medical criterion includes the following items for brain death determination: (i) Coma test, (ii) Pupil test, (iii) Brain stem reflexes test, (iv) Apnea test, (v) EEG confirmatory test.

Generally, risks occur in the standard process of brain death determination. In order to develop a reliable, safety and rapid method in the determination of brain death, the EEG-based preliminary examination has been proposed [2]. To extract informative features from noisy EEG signals and evaluate their significance, several complexity measures are developed for the quantitative EEG analysis in our previous study [2].

In this paper, we present a complexity measures based method to study the recorded EEG signals. Approximate entropy is first applied to evaluate the

K. Yang (✉)
Institute for Cognitive Neurodynamics, East China University of Science and Technology, Shanghai 200237, China
e-mail: yangkun.y@163.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_110,

differences between comatose patients and quasi-brain-deaths. Furthermore, we apply a dynamic complexity measure based on ApEn to EEG data in a time period. Significant statistical difference between the presence and absence of brain activities is obtained. The experimental results tend to show that the proposed method is effective in distinguishing from two groups.

2 EEG Analysis with Complexity Measures

The approximate entropy (ApEn) is a quantity that measures the regularity or predictability of a random signal or time series [3].

To compute the ApEn(m,r) (m: length of the series of vectors, r: tolerance parameter) of a time series $\{x(k)\}$, $(k = 1, \ldots, N)$, $v(k) = [x(k), x(k+1], \ldots, x(x+m-1)]^T$ is first constructed from the signal samples $\{x(k)\}$. Let $D(i, \mathbf{j})$ denote the distance between two vectors v(i) and v(j) $(i, j \leq N - m + 1)$, which is defined as the maximum difference in the scalar components of v(i) and v(j), or $D(i, j) = \max_{1=1,\ldots,m} |v_1(i) - v_1(j)|$. Then, compute the metric $N^{m,r}(i)$, which represents the total number of vectors v(j) whose distance with respect to the generic vector v(i) is less than r, or $D(i, j) \leq r$. Now define $C^{m,r}(i)$, the probability to find a vector that differs from v(i) less than the distance r, as follows:

$$C^{m,r}(\mathbf{i}) = \frac{N^{m,r}(i)}{N - m + 1}, \quad \phi^{m,r} = \frac{\sum_{i=1}^{N-m+1} \log C^{m,r}(i)}{N - m + 1}$$

For m+1, repeat above steps and compute $\phi^{m+1,r}$. ApEn statistic is given by

$$\mathrm{ApEn}(m, r) = \phi^{m,r} - \phi^{m+1,r}.$$

The typical values m=2 and r between 10 and 25% of the standard deviation of the time series $\{x(k)\}$ are often used in practice [3]. As illustrated in Fig. 1, smaller ApEn implies a greater likelihood that similar patterns of measurements will be followed by additional similar measurements like sine wave. If the time series is highly irregular, the occurrence of similar patterns will not be predictive for the following measurements, and ApEn will be relatively large.

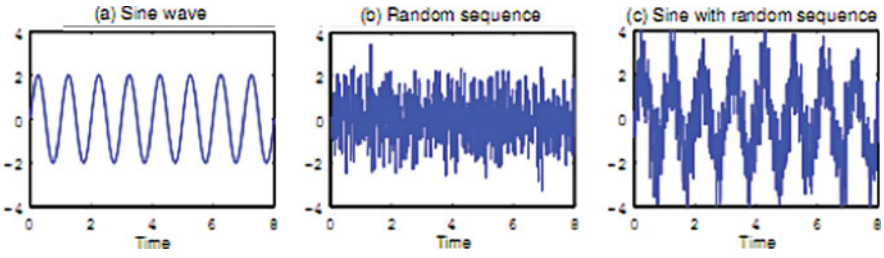


Fig. 1 ApEn of a sine wave, random sequence and sine with random sequence is 0.1834, 0.9362 and 0.5841, respectively

3 EEG Data Analysis with Active Complexity

3.1 EEG Signals and Brain Activity

EEG preliminary examination was carried out in the Shanghai Huashan Hospital affiliated to Fudan University (China). The EEG data were directly recorded at the bedside of the patients by using the portable NEUROSCAN ESI system in ICU. Electrodes (F_{p1}, F_{p2}, F_3, F_4, F_7, F_8, GND, A_1, A_2) of the international 10–20 system are chosen. Sampling rate of EEG was 1,000 Hz and the resistances of the electrodes were less than 8 kΩ.

From June 2004 to March 2006, a total of 35 coma and quasi-brain-death patients had been examined. The patients were classified into a coma group (19 cases) and a quasi-brain-death group (16 cases) after medical diagnosis. One patient behaved from the coma state to quasi-brain-death. On October 18, 2005, the patient was in a deep-coma state. With our proposed signal processing procedure, a theta-wave component was found. However, in the next day, the patient became completely unresponsive to external visual, auditory, and tactile stimuli. No brain wave activity was found. By applying the robust ICA algorithm to the 35 patients' EEG recordings, brain activities δ, θ or α that dominate in lower frequency bands have been extracted from coma cases [2]. The EEG signals of coma cases are brain waves with interfering noises, while the EEG signals of quasi-brain-deaths are interfering noises. Therefore, we consider that regular or predictable components exist in EEG signals of coma. The ApEn measures of coma states may be much lower. Analysis results from the following aspects will be given.

3.2 ApEn of Patients in Coma and Quasi-brain-death States

The EEG recordings usually lasted hundreds of seconds. Generally, when we calculate ApEn measures, the length of a time series is below 5 s. In this part, the length of an EEG signal that we choose is 1 s. We used ApEn measures to analyze all EEG signals that we had recorded. In accordance with the results, we divided them into two different groups: coma and quasi-brain-death. In order to demonstrate the differences between the two types, we present one case from each group, and a patient who behaved from coma to brain death. We use Apen measure to analyze EEG signals of the above patients, and the results are shown in Table 1. The ApEn measure is lower for a coma state than for a quasi-brain-death.

3.3 Dynamic Complexity for Coma and Quasi-brain-death

In this part, along the time-coordinate of EEG signals, we can calculate the ApEn in each second. The window is set on 1,000 samples. Figure 2 is a result of dynamic complexity measures for a coma case, a quasi-brain-death case and a case who behaved from coma to quasi-brain-death. The x-coordinate in Fig. 2 is the time of each case's EEG signal.

Table 1 Summary of ApEn results of two patients in coma state (C) or brain death (D) and a specific patient in different states. Tolerance parameter r equals 0.15 or 0.25

	Two patients				A patient has two states			
	C	D	C	D	C	D	C	D
Chan.	$r = 0.15$		$r = 0.25$		$r = 0.15$		$r = 0.25$	
F_{p1}	0.2609	1.0695	0.4582	1.3366	0.1048	1.1345	0.1978	1.3772
F_{p2}	0.2644	1.1682	0.4459	1.4102	0.1195	1.1041	0.2198	1.3416
F_3	0.1979	0.7289	0.3442	1.0049	0.1207	0.9670	0.2199	1.2600
F_4	0.2627	0.7917	0.4452	1.1277	0.1524	0.9766	0.2620	1.2555
F_7	0.2461	0.8785	0.3853	1.1696	0.2143	1.3082	0.3856	1.4267
F_8	0.2909	1.2872	0.5050	1.3948	0.1631	1.1351	0.2834	1.2662

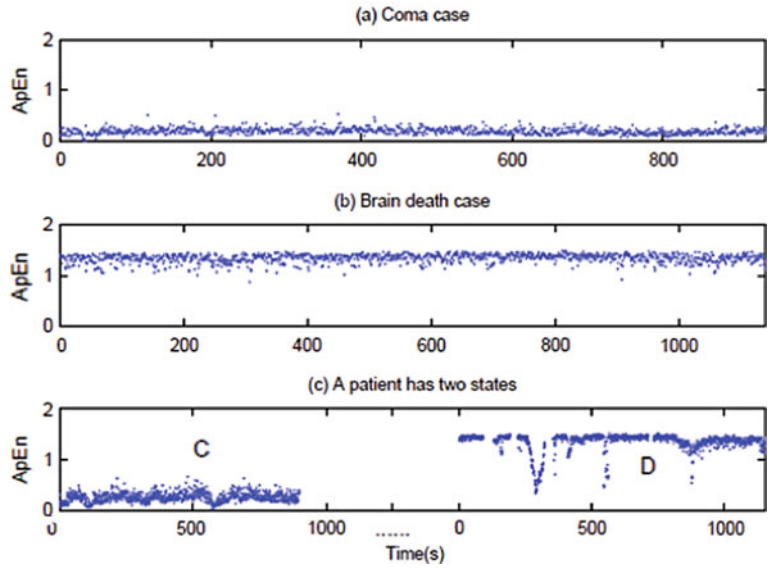


Fig. 2 Dynamic complexity measure for two different cases (**a**) vs. (**b**) and a patient has two states (**c**). r = 0.25

3.4 Evaluation of All Patients' EEG ApEn Results

In our research, all recorded EEG signals of 35 patients are analyzed by ApEn. Except a few patients, we obtained a good classification. In Fig. 3, we give the distribution of statistical ApEn results from the six channels. C represents coma cases while D represents quasi-brain-deaths. The maximum, minimum and average value of ApEn we obtained are displayed. From Fig. 3, the statistical evaluation confirms the existence of regular or predictable components in the EEG signals of coma cases as well as the absence of those in quasi-brain-deaths.

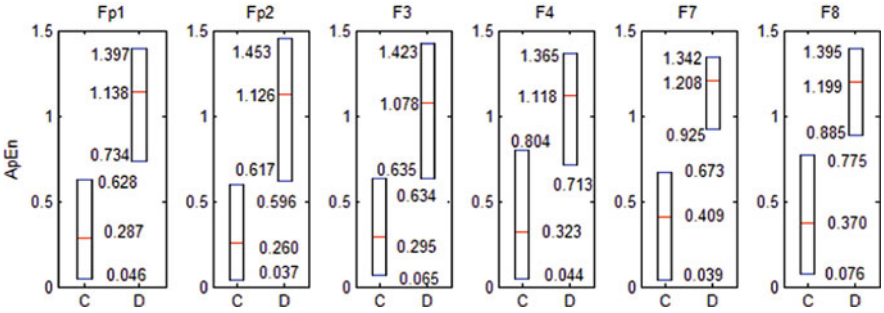


Fig. 3 ApEn statistics (for 6 channels) between coma (C) and quasi-brain-death (D)

4 Conclusion

In this paper, we have applied approximate entropy (ApEn) based dynamic complexity measure to analyze all recorded EEG. For all recorded EEG signals, the above experiments demonstrate the ApEn measures of coma patients and quasi-brain-deaths have different distributions. The dynamic complexity measure, as well as the ApEn measure, is lower for a coma patient than for a quasi-brain-death. Since EEG recording might be corrupted by strong interfering noise, the ApEn measures of some coma patients are high. Although a few cases don't strictly accord with our expectation, we still obtain statistical significance differences between comatose patients and quasi-brain-deaths.

Acknowledgements This work is supported by National Nature Science Foundation of China (10672057, 10872068).

References

1. Wijdicks, E.: Brain death worldwide: accepted fact but no global consensus in diagnostic criteria. Neurology **58** (2002) 20–25.
2. Cao, J., Chen, Z.: Advanced EEG signal processing in brain death diagnosis. In: Signal Processing Techniques for Knowledge Extraction and Information Fusion. Springer (2008) 275–298.
3. Pincus, S.M.: Approximate entropy (ApEn) as a measure of system complexity. Proc. Natl. Acad. Sci. USA **88** (1991) 110–117.

Complexity Analysis of EEG Data with Multiscale Permutation Entropy

Gaoxiang Ouyang, Chuangyin Dang, and Xiaoli Li

Abstract In this study, we propose a powerful tool, called multiscale permutation entropy (MPE), to evaluate the dynamical characteristics of electroencephalogram (EEG) at the duration of epileptic seizure and seizure-free states. Numerical simulation analysis shows that MPE method is able to distinguish between the stochastic noise and deterministic chaotic data. The real EEG data analysis shows that a high entropy value is assigned to seizure-free EEG recordings and a low entropy value is assigned to seizure EEG recordings at the major scales. This result means that EEG signals are more complex in the seizure-free state than in the seizure state.

Keywords Multiscale permutation entropy · Epileptic EEG · Complexity

1 Introduction

The electroencephalogram (EEG) signal is a measure of the summed activity of approximately 1–100 million neurons lying in the vicinity of the recording electrode [1], and often exhibits complex fluctuations that are not simply due to "contaminative" noise but contain information about the underlying dynamics. Traditional entropy algorithms are single-scale based and, therefore fail to account for the multiple scales inherent in brain electrical activities [2]. A new method, called multiscale entropy (MSE), has recently been proposed by Costa et al. [2] to measure the complexity of a time series. This complexity method is to consider the correlations over multiple scales of a time series instead of a single scale. In this MSE method, sample entropy (SampEn) is applied for different scales. SampEn is a measure of regularity (orderliness) in time series; two parameters: the sequence length m and the tolerance level r are needed to be determined in advance. SampEn increases with the decrease of *r value* because the criterion for sequence matching becomes more stringent [3].

G. Ouyang (✉)
Department of MEEM, City University of Hong Kong, Kowloon, Hong Kong
e-mail: gx.ouyang@gmail.com

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_111,

In general, there are two approaches to select the parameter r in MSE method: (i) r is fixed for all scales; (ii) r is calculated for each scale separately. Therefore, the outcome of the MSE algorithm will change with the selection of parameter r [4].

Recently, Bandt and Pompe introduced permutation entropy as a new measure for the irregularity of a non-stationary time series, which does not require a reconstruction of an attractor in state space [2]. This method is to consider the order relations between time series as a result, and it is robust in the presence of observational and dynamical noise [5]. Comparing with the SampEn, the calculation of the permutation entropy is very simple and fast [6]. These advantages of the permutation entropy motivate us to explore whether or not the permutation entropy can replace SampPE to estimate the multiscale entropy of EEG recordings; this method is denoted as multiscale permutation entropy (MPE).

2 Materials and Methods

2.1 Animal Experiments and EEG Series

All procedures were performed under a British Home Office project license (UK Animals (Scientific Procedures) Act, 1986). Experiments were performed in 24 male Genetic Absence Epilepsy Rats from Strasbourg (GAERS) of at least 13 weeks of age. Much of the details on the experiments can be found in [7]. The signal from the EEG electrode was directly visualized on an oscilloscope and was further amplified (BioAmp ML 136), filtered, digitized (100 Hz) and stored using a PowerLab 2/20 running Chart v4.2 software (ADInstruments, Hastings, UK). The EEG data sets were preprocessed by a band-pass filter at 0.5–22 Hz. To investigate the complexity of EEG data during seizure state, EEG signals of epileptic rats were selected and dissected from seizure-free and seizure states, each containing 50 single-channel EEG signals of 20 s duration.

2.2 Multiscale Permutation Entropy (MPE)

MPE method, similar with MSE method, incorporates two procedures. First, a "coarse-graining" process is applied to a time series. For a given time series $\{x_1, x_2, \cdots, x_N\}$, construct a consecutive coarse-grained time series by averaging a successively increasing number of data points in non-overlapping windows. Each element of a multiple coarse-grained time series $y_j^{(\tau)}$ is calculated according to the following equation

$$y_j^{(\tau)} = 1/\tau \sum_{j=(j-1)\tau+1}^{j\tau} x_i \,. \tag{1}$$

where τ represents the scale factor and $1 \leq j \leq N/\tau$. The length of each coarse-grained time series is the integral part of N/τ.

Next, the permutation entropy [5] is calculated for each coarse-grained time series, and then plotted as a function of the scale factor τ. To compute the permutation of a coarse-grained time series y_j, first the series of vectors of length m, $v^m(n) = [y_n, y_{n+1}, \cdots, y_{(n+m-1)}]$, $1 \leq n \leq N/\tau - m + 1$ is derived from the signal samples y_j. Then, $v^m(n)$ can be arranged in an increasing order: $[y_{n+j_1-1} \leq y_{n+j_2-1} \cdots \leq y_{n+j_n-1}]$. For m different numbers, there will be $m!$ possible order patterns π which are also called permutations. For a permutation with number π, let $f(\pi)$ denote its frequency in the time series. Then the relative frequency is $p(\pi) = f(\pi)/(N/\tau - m + 1)$. The permutation entropy for the time series is defined as

$$H(m) = -\sum_{\pi=1}^{m!} p(\pi) \ln p(\pi). \tag{2}$$

The largest value of $H(m)$ is $\log(m!)$, which means all permutations have equal probability. The smallest value of $H(m)$ is zero, which means the time series is very regular. In other words, the smaller the value of $H(m)$, the more regular the time series is. Permutation entropy refers to the local order structure of the time series, which can give a quantitative complexity measure for a dynamical time series. Permutation entropy calculation only depends on the selection of m. When m is too small (less than 3), the scheme will not work well, since there are only very few distinct states for EEG recordings. Often, for a long EEG recording a large value of m is better. In this study, we calculate for all EEG data with the parameters $m = 4$.

3 Results

3.1 Numerical Simulation Analysis of Chaotic System

To demonstrate the performance of this method, MPE method is firstly applied to analyze simulated stochastic noise and deterministic chaotic data. We test the MPE method on simulated white noise, 1/*f* noises and fractional Brownian noise (with Hurst exponent 0.5), as well as on the synthetic chaotic data sets: logistic map, Henon's map, Lorenz map and Rossler's oscillator [8]. As can be seen in Fig. 1a, b, the entropy measures for the deterministic chaotic data increase on small scales and then trend to relatively constant value; the variation of the permutation entropy for all chaotic data sets has a similar behavior. Moreover, this trend of the variation of the permutation entropy with scale is different from the stochastic noise data. The MPE measures for noise data decrease on small scales and then trend to relatively constant value. This result is consistent with the fact that, unlike stochastic noise, deterministic chaotic systems contain complex structures across multiple time scales. It is suggested that the MPE analysis can be used to distinguish between the stochastic noise and deterministic chaotic data.

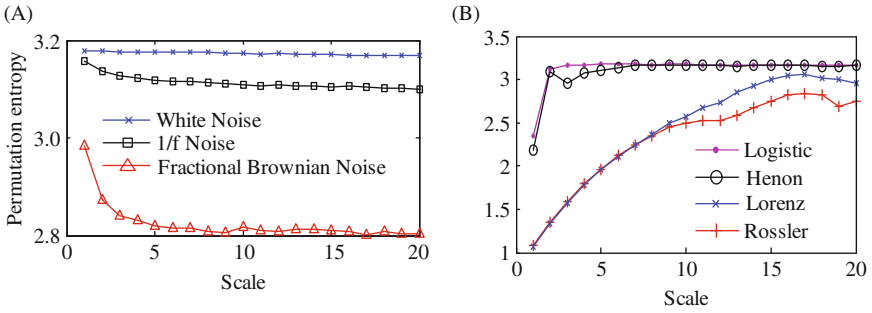


Fig. 1 The result of the MPE analysis on simulated noise date (**a**) and chaotic data sets (**b**). *Symbols* represent results of simulations for time series of 4,000 points

3.2 MPE Analysis of EEG

Next, MPE method is applied to analyze the EEG recordings. Permutation entropy is evaluated for 20 scale factors with the dimension m=4. The result of the MPE analysis on all EEG data is shown in Fig. 2. For scale 1, which is the only scale considered by traditional single-scale based method, the permutation entropy values for EEG segments are averaged at 2.44±0.04 in seizure-free state and 2.34±0.10 in seizure state, respectively. The entropy values in seizure-free are significantly larger than those in seizure state for most scales ($P < 0.05$), which are consistent with the hypothesis of loss of complexity with epileptic seizure [9]. The entropy measures for EEG increase on small scales, then trend to relatively constant value for larger scales. The MPE curve firstly reaches to constant value about at a scale factor of 3–4 in seizure-free state. However, the MPE curve for the seizure EEG data has common features: it has a local maximum entropy value at a scale factor of 3–4; then it gradually decreases and reaches to local minimum entropy value at a scale factor of 6–7; after reaching the minimum entropy, it gradually increases and trends to relatively constant value for larger scales. These patterns of MPE curves are distinctly different between the seizure-free and seizure state.

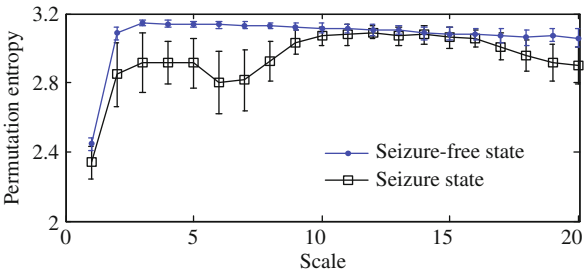


Fig. 2 MPE analysis of EEG recordings in seizure-free state and seizure state. *Symbols* represent the mean values of entropy for each group and bars represent the standard deviation

4 Conclusions

In this study, MPE method is proposed for analyzing complex time series, including simulated stochastic noise and deterministic chaotic data, as well as real data. The entropy measure for the deterministic chaotic data increases on small scales and then stabilizes to relatively constant value. However, the MPE measure for noise data decreases on small scales and then stabilizes to relatively constant value. This result is consistent with the fact that, unlike stochastic noise, deterministic chaotic systems contain complex structures across multiple time scales. Therefore, the MPE method can be used to distinguish between the stochastic noise and deterministic chaotic data.

On the other hand, the entropy measure for EEG increases on small scales and then stabilizes to relatively constant value for larger scales. The MPE curve firstly reach to constant value about at a scale factor of 3–4 in seizure-free state. During the seizure state, the MPE curve has a local minimum entropy value at a scale factor of 6–7. The variation of the permutation entropy for EEG data and chaotic data sets showed a similar behavior. Moreover, the entropy values in seizure-free are significantly larger than those in seizure state for majority of scales ($P < 0.05$). This is because epilepsy may relate to their increased regularity (predictability) associated with the synchronous discharge of large numbers neurons [7]. These results suggest that the MPE method might be a powerful tool to reveal the hidden characteristics of the epileptic EEG signals.

Acknowledgments This research was supported by Strategic Research Grant of City University of Hong Kong (7002224) and Program for New Century Excellent Talents in University (NECT-07-0735).

References

1. Sleigh, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L., Grant, C., Ludbrook, G.: Cortical entropy changes with general anaesthesia: theory and experiment. Physiological Measurement **25** (2004) 921–934.
2. Costa, M., Goldberger A.L., Peng C.K.: Multiscale entropy analysis of physiologic time series. Phys. Rev. Lett. **89** (2002) 062102 .
3. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circ. Physiol. **278** (2000) 2039–2049.
4. Nikulin, V.V., Brismar, T.: *Comment on "multiscale entropy analysis of* complex physiological time series". Phys. Rev. Lett. **92** (2004) 089803.
5. Bandt, C., Pompe, B.: Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. **88** (2002) 174102.
6. Bahraminasab, A., Ghasemi, F., Stefanovska, A., McClintock, P.V.E., Kantz, H.: Direction of coupling from phases of interacting oscillators: A permutation information approach. Phys. Rev. Lett. **100** (2008) 084101.
7. Li, X., Ouyang, G., Richards, D.A.: Predictability analysis of absence seizures with permutation entropy. Epilepsy Res. **77** (2007) 70–74.
8. Rosso, O.A., Larrondo, H.A., Martin, M.T., Plastino, A., Fuentes, M.A.: Distinguishing noise from chaos. Phys. Rev. Lett. 99 (2007) 154102.
9. Ouyang, G., Li, X., Dang, C., Richards, D.A.: Using recurrence plot for determinism analysis of EEG recordings in genetic absence epilepsy rats. Clin. Neurophysiol. **119** (2008) 1747–1755.

Cueing-Dependent Activity in Bilateral M1 in Parkinson's Disease

Andreas Daffertshofer, Alistair N. Vardy, and Erwin E.H. van Wegen

Abstract We investigated rhythmic force production during two different types of cueing; regular and irregular sequences of tactile stimuli. We expected the corresponding movement-related cortical activity to be most distinguishable in the beta frequency range (13–30 Hz). This accompanying beta activity was expected to show a differential effect of cueing. Twenty Parkinson patients and fifteen age-matched healthy controls performed 3×30 s series of rhythmic movements separated by rest. Each series was performed using (ir-)regular cueing. Using linear beamformers, MEG recordings revealed dominant motor-related activity in bilateral M1s. As expected, at theses sources the Hilbert amplitude of alpha and beta oscillations displayed reduced activity during motor performance. We found a sustained beta rebound during subsequent resting states to a higher level compared to pre-performance. This rebound lasted 5–10 s in both Parkinson patients and controls and its strength depended on cueing type. We also found a likewise sustained effect in the alpha band, where amplitude remained low after movement termination. Although quality of performance largely agreed for patients and controls, dependent on cueing type both the beta rebound and the alpha after-effect differed. The latter discriminated between groups which indicates that Parkinson patients used the same strategy (or attentional load) for mere rhythmic performance (regular cue) or tracking (irregular cue).

Keywords Parkinson's disease · Motor performance · MEG activity · Spectral power

1 Introduction

Most prominent of all dysfunctions in Parkinson's disease (PD) are movement-related problems [1], which are suspected to originate in a disturbed functioning

A. Daffertshofer (✉)
Research Institute MOVE, VU University, Amsterdam, The Netherlands
e-mail: a.daffertshofer@fbw.vu.nl

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1_112,

of the neural circuitry between thalamus, basal ganglia, and primary motor cortices [2]. In particular, the timing of movements becomes problematic in PD as patients have difficulties initiating, continuing, and terminating movements alike. In view of its profound impact on daily functioning, maintaining motor timing has hence received quite some attention. External *cueing* has proven very beneficial in PD. For example, when provided with rhythmic auditory or visual cues, the shuffled, irregular gait in PD becomes regular and is easier maintained [3–5]. What aspects of cueing make it beneficial? Cueing provides a patient with a stimulus adding a temporal reference or a trigger. Is the presence of a cue sufficient and is its regularity essential? To answer this, we compared effects of regular cueing with sequences of cues that are not equidistance but have a small but noticeable jitter. We hypothesized that the rhythmicity of cues is its key-ingredient for improvement of motor timing. When comparing PD patients to healthy controls we therefore expected to find the least differences in performance and its neural correlates by means of magnetoencephalography (MEG) in the context of rhythmic cues.

2 Methods

Twenty PD patients (PD, five female) and fifteen healthy age-matched controls (CO, three female) participated in this study. PD patients were recruited from existing patient databases from the RESCUE project (www.rescueproject.org) and the outpatient clinic for movement disorders of the VU University medical center. PD patients were 62.3±2.5 years of age, had an Minimal Mental State Examination (MMSE) of 29.2±1.5, Hoeh and Yahr stage (HY) of 2.6±0.5, and a PD duration of 6.3±3.5 years (range refers to mean ±SD).

2.1 Procedure and Data Acquisition

Lying supine with eyes open, subjects received written instructions about conditions projected onto the ceiling of the magnetically shielded room housing the MEG system. Subjects were asked to perform a rhythmic motor task with their right hand consisting of squeezing an air-filled rubber bulb. The pace for the task was set by a sensory stimulation delivered by an expandable membrane attached to the ventral side of the left index finger delivering a 200 ms pulse at 200 kPa.

In a nutshell, we investigated two cued conditions, with regular and irregular cues (RC and IC, respectively) and two conditions in which subjects performed self-paced movements (not reported here) framed by extended periods of rest during which MEG was also recorded; Fig. 1 for a sketch. Each condition was presented once adding up to a total of a measurement time of 34.5 min. The order of the conditions was counterbalanced over subjects.

Brain activity was recorded using a 151-channel whole-head MEG system (CTF Systems Inc., Vancouver, Canada) using 3rd-order synthetic gradiometers. Signals were low-pass filtered at 200 Hz prior to digitization at a rate of 625 Hz.

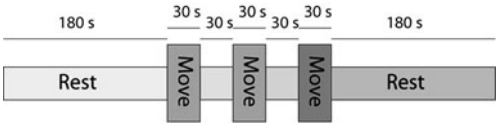


Fig. 1 Experimental protocol: all conditions started with 180 s of rest followed by three 30 s movement periods consisting of cued squeezing separated by 30 s rest periods, and concluded by another 180 s of rest

2.2 Data Analysis

Peaks in the pressure signals were determined to quantify performance: the total error of synchrony was defined as the sum of the absolute difference between the moments of maximal pressure and the stimulus times.

2.2.1 Source reconstruction

MEG signals were transformed to source space via synthetic aperture magnetometry (SAM) beamformers [6] using the average of 152 T1-weighted stereotaxic volumes from the ICBM project (www.bic.mni.mcgill.ca). Task-relevant areas in the beta band are shown in Fig. 2: maximal significant contrasts between power during 3×20 s of movement and 3×20 s rest were found in the bilateral M1s onto which the MEG were projected yielding the time-dependent activity in $M1_{left/right}$.

2.2.2 Time-frequency analysis

$M1_{left/right}$ activities were assessed via the power in different time-windows within distinct frequency bands; Fig. 3. We selected alpha and beta band (7–11 and 13–30 Hz, respectively; we note that the gamma band, 40–70 Hz, did not yield any

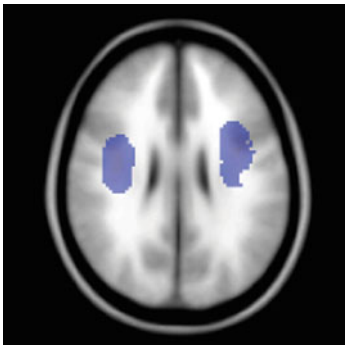


Fig. 2 SAM analysis [7]: we compared beta band power during 3×20s of movement with 3×20 s rest (each 20 s interval centered in the corresponding 30 s block to avoid transients; cf. Fig. 1) for every voxel and subject. The resulting pseudo-*t* values were averaged over subjects for each group and condition yielding maximal significant contrasts

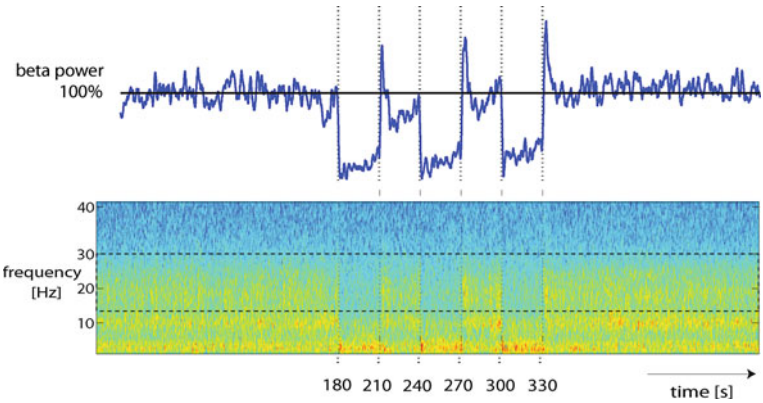


Fig. 3 Time-frequency analysis: after initial short-time Fourier analysis, data were filtered in different frequency bands prior to computing the time-dependent Hilbert amplitude (*upper panel* shows the amplitude in the beta band). Per frequency band signals were normalized by dividing by the mean amplitude over 120 s of the pre-movement rest period, i.e. these values were set to 100%). As to-be-assessed time-windows we first used the aforementioned 20 s intervals centered within the three 30 s movement period and averaged the amplitude.

significant difference between groups and conditions). The resulting two values per subject, condition, and frequency band (alpha and beta) were evaluated statistically. Here we only report the test for effects in the post-movement period by computing the mean amplitude over a 5 s interval of rest 2 s after the final movement period.

2.3 Statistics

Performance was assessed using a 2-way mixed design ANOVA (between-subject variable *group*: PD patients or controls (PD or CO); and within-subjects variable *cueing type*: RC or IC). Comparison between pre- and post-movement resting states MEG signals was assessed with a $2\times2\times2$ mixed design ANOVA (between-subject variable *group*: PD or CO; within-subjects variables *source*: $\mathrm{M1}_{\mathrm{left}}$ and $\mathrm{M1}_{\mathrm{right}}$, and *cueing type*: RC or IC) for each frequency band separately. A significance level of $\alpha = 0.05$ was used.

3 Results

There was a significant difference in motor performance between the RC and IC condition, where the error was larger for the IC condition ($F_{1,15} = 142.8, p < 0.001$). Performance was similar for CO and PD.

For the pre- and post-movement power changes, both the alpha and beta frequency bands showed a significant difference. The latter showed a significant increase in power in the first 5 s after the final movement termination; Fig. 4. This

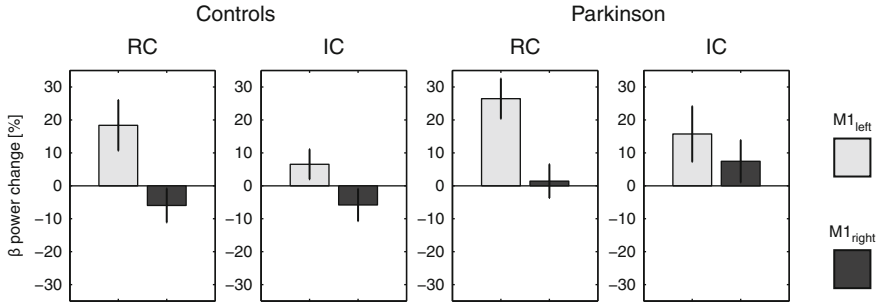


Fig. 4 After-effects in the beta band: we found a main effect of *source* ($F_{1,25} = 20.78, p < 0.001$): changes for $M1_{left}$ were more positive and larger than those for $M1_{right}$. We also found a *cueing type* × *source* interaction ($F_{1,25} = 7.27, p = 0.012$): for $M1_{left}$, the changes were larger during RC than for IC; also, for $M1_{right}$ during RC the changes were negative and during IC the changes were positive

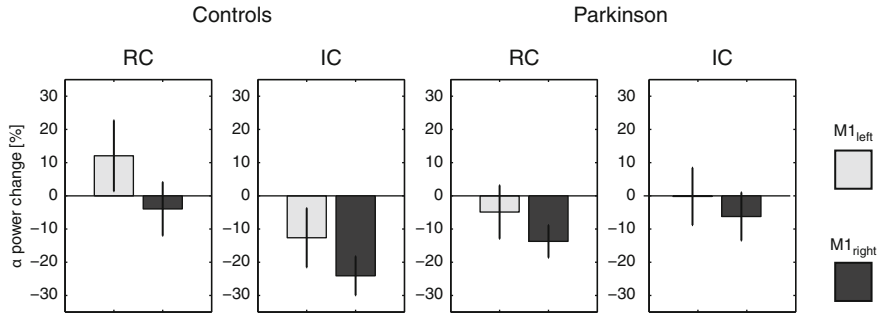


Fig. 5 After-effects in the alpha band: we found a main effect of *source* ($F_{1,25} = 7.28, p = 0.012$): the changes for $M1_{left}$ were more positive than for $M1_{right}$. There was also a *cueing type* × *group* interaction ($F_{1,25} = 4.65, p = 0.041$): changes during RC were positive for CO but negative for PD; during IC, differences for CO were negative whereas there was no such change in sign for PD

effect largely matched for CO and PD. In the alpha band amplitude dropped after the final movement block as compared to the pre-movement resting state. This differed between RC and IC conditions for CO but not for PD; Fig. 5.

4 Conclusion

Rhythmic cueing can elicit differential responses in the cortical activity in early stage PD. We localized the activity in bilateral M1 during an experimental paradigm involving rhythmic unimanual force production paced by regular and irregular cues. Task performance matched between PD patients and healthy controls. M1 activity in the alpha and beta frequency bands, however, differed in particular during an up-to 5 s lasting resting period after movement.

Following the irregular cue may require more attention because of its unpredictability and this increased attention may cause the performance terminating beta rebound to disappear. Whether that difference in effect is also reflected in other parts of the neural axis remains to be seen; we plan to assess the cortico-muscular synchrony in further detail. Strikingly, however, in the alpha band the effect of cueing type was absent in Parkinson patients, therefore the changes in attention might have been be less pronounced while performance remained comparable to that of the healthy subjects. Put differently, in order to achieve proper performance, PD patients used the same strategy (or attentional load) for both rhythmic performance (regular cue) or tracking (irregular cue).

Acknowledgments Andreas Daffertshofer: Thanks to the Netherlands Foundation for Scientific Research (grant #452-04-344). Erwin E.H. Van Wegen: Thanks to the Stichting Internationaal Parkinsonfonds (project # IPF-2006-1).

References

1. Wolters, E., van Laar, T., and Berendse, H.: Parkinsonism and Related Disorders. Amsterdam VU University Press (2007).
2. DeLong, M.R. Wichmann, T.: Circuits and circuit disorders of the basal ganglia. Arch. Neurol, 64(1) (2007): 20–4.
3. McIntosh, G. C., Brown, S. H., Rice, R. R., and Thaut, M. H. Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. J. Neurol. Neurosurg. Psychiatry. 62(1) (1997): 22–6.
4. van Wegen, E., de Goede, C., Lim, I., Rietberg, M., Nieuwboer, A., Willems, A., Jones, D., Rochester, L., Hetherington, V., Berendse, H., Zijlmans, J., Wolters, E., and Kwakkel, G.: The effect of rhythmic somatosensory cueing on gait in patients with Parkinson's disease. J. Neurol. Sci. 248(1–2) (2006): 210–4.
5. Nieuwboer, A., Kwakkel, G., Rochester, L., Jones, D., van Wegen, E., Willems, A. M., Chavret, F., Hetherington, V., Baker, K., and Lim, I.: Cueing training in the home improves gait-related mobility in Parkinson's disease: the RESCUE trial. J. Neurol. Neurosurg. Psyc. 78(2) (2007): 134–140.
6. Vrba, J. and Robinson, S. E.: Signal processing in magnetoencephalography. Methods 25(2) (2001): 249–271.
7. Nichols, T. E. and Holmes, A. P.: Nonparametric permutation tests for functional neuroimaging. Hum. Brain. Mapp. 15(1) (2002): 1–25.

Index

R. Wang, F. Gu (eds.), *Advances in Cognitive Neurodynamics (II)*,
DOI 10.1007/978-90-481-9695-1,