A Recurrence-Based Approach for Feature Extraction in Brain-Computer Interface Systems

Luisa F. S. Uribe, Filipe I. Fazanaro, Gabriela Castellano, Ricardo Suyama, Romis Attux, Eleri Cardozo and Diogo C. Soriano

Abstract The feature extraction stage is one of the main tasks underlying pattern recognition, and, is particularly important for designing Brain-Computer Interfaces (BCIs), i.e. structures capable of mapping brain signals in commands for external devices. Within one of the most used BCIs paradigms, that based on Steady State Visual Evoked Potentials (SSVEP), such task is classically performed in the spectral domain, albeit it does not necessarily provide the best achievable performance. The aim of this work is to use recurrence-based measures in an attempt to improve the classification performance obtained with a classical spectral approaches for a five-command SSVEP-BCI system. For both recurrence and spectral spaces, features

L.F.S. Uribe · R. Attux · E. Cardozo
Department of Computer Engineering and Industrial Automation (DCA/FEEC),
University of Campinas (UNICAMP), Av. Albert Einstein, 400,
Campinas, SP 13083-852, Brazil
e-mail: lsuarez@dca.fee.unicamp.br

R. Attux
e-mail: romis@dca.fee.unicamp.br

E. Cardozo
e-mail: eleri@dca.fee.unicamp.br

G. Castellano
Neurophysics Group, Deptartment of Cosmic Rays and Chronology,
Institute of Physics Gleb Wataghin (IFGW), UNICAMP, Rua Sérgio Buarque de Holanda,
777, Campinas, SP 13083-859, Brazil
e-mail: gabriela@ifi.unicamp.br

F.I. Fazanaro · R. Suyama · D.C. Soriano (✉)
Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas (CECS),
Universidade Federal do ABC, Avenida dos Estados, 5001, Santo André, SP 09210-580, Brazil
e-mail: diogo.soriano@ufabc.edu.br

F.I. Fazanaro
e-mail: filipe.fazanaro@ufabc.edu.br

R. Suyama
e-mail: ricardo.suyama@ufabc.edu.br

N. Marwan et al. (eds.), *Translational Recurrences*, Springer Proceedings
in Mathematics & Statistics 103, DOI 10.1007/978-3-319-09531-8_6

were selected using a cluster measure defined by the Davies-Bouldin index and the classification stage was based on linear discriminant analysis. As the main result, it was found that the threshold ε of the recurrence plot, chosen so as to yield a recurrence rate of 2.5 %, defined the key discriminant feature, typically providing a mean classification error of less than 2 % when information from 4 electrodes was used. Such classification performance was significantly better than that attained using spectral features, which strongly indicates that RQA is an efficient feature extraction technique for BCI.

1 Introduction

The main objective of a Brain-Computer Interface (BCI) is to provide an alternative communication channel for human beings without using the classical biological efferent pathways. In the last decade, these systems have experienced a remarkable development due to the greater availability of low cost instrumentation and computational resources [1, 2]. In particular, the use of BCI in the context of assistive technology is very important for people suffering from stroke, spinal cord injury, degenerative disorders (e.g. amyotrophic lateral sclerosis), and any conditions that impose drastic limitations to communication and mobility [3]. Presently, it is estimated that, only in Europe, there are almost 300 thousand people with spinal cord injury (SCI), with eleven thousand new injuries occurring every year [4]. Moreover, approximately 40 % of the total population of patients with SCI are quadriplegics and it is known that loss of motor functions significantly decreases the quality of life [2].

To accomplish a rehabilitation purpose, BCI systems typically consist of well-characterized components, being signal acquisition, processing and feedback stages the main ones, as illustrated in Fig. 1. Within this general scheme, the typical approach for brain signal acquisition relies on the recording of surface EEG, being the processing and feedback stages strongly dependent on the adopted BCI paradigm [1]. The BCI paradigm refers to the process of inducing a stable electrical pattern in the brain to be detected and recognized, being finally associated with an external command. A classical example is that of task imagery, i.e. the request to perform a mental task associated with well-defined motor actions, something that can be detected in a specific cortical area (e.g. the motor cortex) and, in general, suitably classified. Another commonly employed BCI paradigm can be formulated in terms of evoked potentials, such as those present in the P300 response—a potential elicited by the process of focusing attention in events occurring with low probability—and the Steady State Visual Evoked Potentials (SSVEP) [1].

In the SSVEP paradigm, different commands (like those defining the direction of the movement of a wheel-chair) are presented to the user as visual stimuli flickering with different frequencies, and the user is requested to focus on the command that he/she wants to perform. In this case, the electrical activity of the visual cortex tends

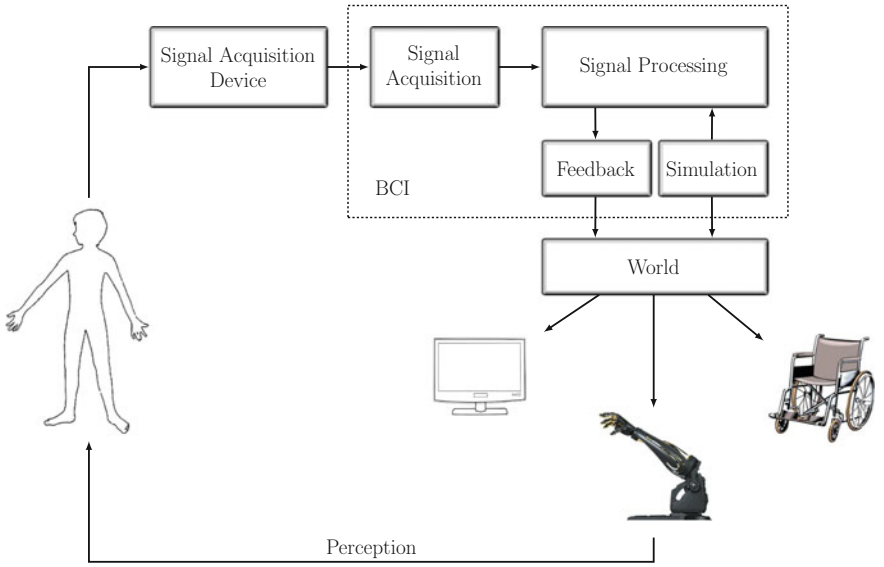

Fig. 1 A general BCI system

to synchronize with the visual stimulus chosen by the user and a simple spectral analysis can be used in order to identify the command selected by the person [5].

Once the BCI paradigm has been defined, the processing and feedback can also be specified in a more precise manner. Most works use a structure composed of feature extraction, feature selection and classification stages. These stages are essential for performing pattern recognition and classification, being thus crucial to define the external command.

Although classical spectral analysis is widely used for feature extraction in SSVEP-BCI systems, it can lead to non-optimal classification results depending on the employed visual stimulation apparatus or require a large amount of data for suitably training the classifier. In this chapter, a new recurrence-based feature extraction approach for BCI systems based on the SSVEP paradigm is presented. It is shown here that the threshold ε of the Recurrence Plot (RP), defined so as to match a specific Recurrence Rate (RR), is the key feature to separate the classes (which correspond to the commands), providing a better performance in comparison to a strategy based on spectral analysis. The results are indicative of the potential of recurrence analysis in the context of BCI, raising quite interesting application perspectives.

This work is organized as follows: Sect.2 presents the instrumentation and the experimental procedure employed in the context of the BCI system used here. This section also briefly introduces the signal processing techniques explored for feature extraction, feature selection and classification. Section 3 exhibits the system performance (in terms of mean classification error) when the proposed recurrence-based

approach is employed and a comparison with different feature selection heuristics operating in the spectral domain. Finally, this chapter is concluded with a discussion about the developed ideas and potential extensions.

2 Methods

2.1 Experimental Setup and Experimental Procedure

In the performed experiments, the brain signals were acquired by means of surface EEG [1], due to its simplicity and non-invasive character. For the acquisition, digitization and amplification of the EEG signals, the g.USBamp system of the g.tec company[1] was employed. This device allows the simultaneous capture of 16 channels, with 24-bit resolution each, through a USB 2.0 connection to a desktop computer. The device is shown in Fig. 2a.

The EEG was recorded using 16 dry electrodes with 8-pin each (the g.SAHARA g.tec system), which are built using a special gold alloy (see Fig. 2a). These electrodes were placed in 16 locations defined by the international 10–20 system [1], which were chosen in accordance with the SSVEP response (i.e. emphasizing signals in the visual cortex): Fz, Cz, Pz, Oz, PO3, PO4, O1, O2, P3, P4, Iz, POz, PO7, PO8, O9, O10. Furthermore, two reference electrodes were placed in each mastoid. Figure 2b

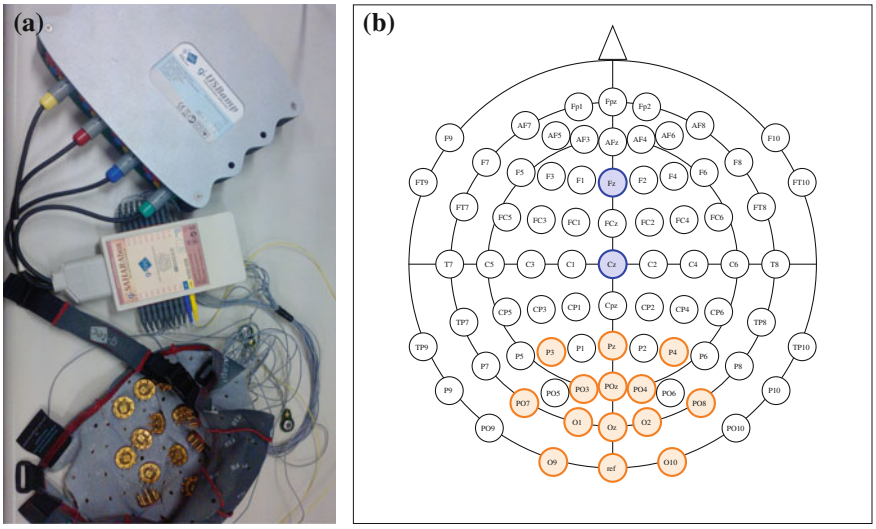

Fig. 2 **a** Equipment and configuration for EEG signals capture. In **b** the *orange circles* represent the visual cortex region and the *blue circles* define other interesting positions along the medial longitudinal fissure

[1] http://www.gtec.at/Products/Hardware-and-Accessories/g.USBamp-Specs-Features.

illustrates the cap layout, in which the letters F, T, C, P and O correspond, respectively, to the frontal, temporal, central, parietal and occipital lobes, with odd numbers used to reference the left hemisphere and even numbers the right one.

Before starting the acquisition session, it was ensured that all electrodes had impedance values below 5 kΩ, which represents a common limit employed during EEG recordings. The amplifier was configured with a bandpass filter of 0.1–60 Hz and a notch filter at 60 Hz for all EEG channels, in order to cancel out DC components interference and noise. A sampling frequency of 128 Hz was used throughout all acquisitions.

In order to register the signals, the BCI2000 software[2] was employed. This software is particularly interesting due to its ability to create different experiments with multiple settings for capturing the data, which were first stored in files with a specific BCI2000 format and subsequently imported by MATLAB, in which functions from the statistical signal processing toolbox were used.

To carry out the visual stimulation, a set of four light-emitting diodes (white LEDs) was used, driven in four different frequencies, namely 13, 18, 21 and 25 Hz. These values were chosen in view of previous studies that showed strong SSVEP responses evoked in this range [5].

Three healthy subjects (ages 21, 25, 28; one woman) with no previous history of neurological diseases participated in this study. The study was approved by the ethics committee of University of Campinas, and all subjects signed an informed consent previous to data acquisition. Subjects sat on a comfortable chair placed at a distance of 0.5 m both from the LEDs and the computer screen. The same ambient light level was maintained during the execution of all experiments. Figure 3a, b show, respectively, the LED set-up and the interface used in the BCI2000 software for the execution of the experiments and for the capture of data.

For every subject, four different sessions for data acquisition were held five minutes apart from each other. The sessions comprised four separate runs interleaved by a pause of one minute. The last session had only three runs.

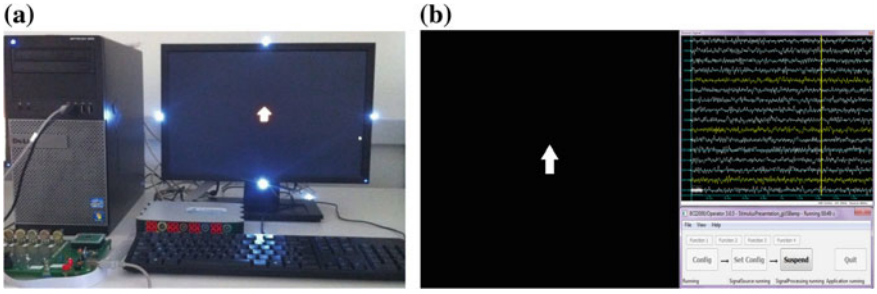

Fig. 3 **a** Visual stimulation platform used to evoke the SSVEP response. **b** Interface requesting the user to focus on a specific LED

[2] http://www.schalklab.org/research/bci2000.

Each run was composed of 25 trials associated with 5 different tasks (four visual stimuli plus no command). Therefore, there were a total of $15 \times 25 = 375$ trials per subject, with $375/5 = 75$ the total number of trials for each task. Subjects were instructed to stay still during every run.

Each image was continuously displayed during 6 seconds. After this period of time, the screen went blank for 1.5 s before the next image appeared. During the resting condition, the subject could rest the eyes. The pause period between each run was long enough to allow the subject to stretch and move around.

2.2 Signal Processing Techniques

After the EEG signals were acquired, a pre-processing stage took place, consisting in the normalization of each recorded time series with respect to the maximum absolute value found therein. Thereafter, feature extraction was performed with the aim of representing each trial of a given electrode in a suitable space, which, in this work, may correspond to either the recurrence or the spectral domain.

The recurrence domain is defined by Recurrence Quantification Analysis (RQA) measures obtained for a given trial. As described by [6], the classical RQA measures correspond to the percentage of determinism (DET), the entropy of the diagonals (ENTR) and the length of the longest diagonal of the map (LMAX) (excluding the main diagonal). The L-∞ norm was used for quantifying the distance between the points in the reconstructed state space, given its lower computational cost [6]. An adaptive recurrence plot was used to minimize the recurrence plot variability, thus avoiding inappropriate choices of ε and providing a better comparison between maps from different EEG electrodes. The RP parameters were defined after a preliminary analysis concerning the use of a variable threshold as reported in [7, 8]. The threshold ε of the map was defined to match a Recurrence Rate (RR) of 2.5 %—with an embedding dimension $m = 5$ and a time delay $\tau = 5$—being also employed as a feature. The DET measure was calculated for different diagonal intervals, being DET1 associated with the percentage of determinism related to diagonal lengths from 5 to 10, DET2 from 5 to 15, DET3 from 10 to 15, DET4 from 10 to 20, DET5 from 15 to 20, DET6 from 15 to 25. This partition is useful to detect deterministic sources with different characteristics [9]. In general, the diagonals obtained in the recurrence plot were small, which justifies the upper limit established by DET6.

The spectral attributes were determined using the classical Welch method for computing the Power Density Spectrum (PDS). In this approach, each trial was divided into 8 sub-blocks of identical length (with 170 points each) using a Hamming window function with 50 % superposition, i.e. with an amount of 85 overlapping points. Each block was once again divided in segments of 128 points: padding with zeros the last segment and wrapping them, the Discrete Fourier Transform (DFT) of each of the eight segments was obtained—implying in a spectral resolution of 1 Hz—and the square of its absolute value calculated, followed by the average concerning all sub-blocks. For implementation details, see [10, 11].

The feature selection stage was implemented based on the cluster measure defined by the Davies-Bouldin (DB) index [12]. This measure combines in a single expression two main relevant aspects of data clustering: the minimization of the intraclass distance and the maximization of the distance between classes, which can be mathematically described by:

$$\mathrm{DB} = \frac{1}{M}\sum_{i=1}^{M}\max_{i\neq j}\left[\frac{\mathrm{diam}\,(C_i)+\mathrm{diam}\left(C_j\right)}{d\left(\mu_i,\mu_j\right)}\right] \tag{1}$$

where $d\left(\mu_i,\mu_j\right)$ is the distance between the centers of classes i and j, $\mathrm{diam}\,(C_i)$ is the maximum distance between all pairs of samples in class i, and M is the number of classes. Hence, low values of the DB index indicate good class discrimination, while higher values indicate less favorable scenarios. For this reason, the inverse of the DB index was employed as a rank measure (DBINV), being each class defined by the set of trials labeled with the same visual stimulus.

Finally, the classification stage was based on Linear Discriminant Analysis (LDA) [13], since it is straightforward to implement, fast to train and widespread in the BCI literature (see, for instance, [14]). In this approach, a linear combination $\mathbf{w}$ of the features $\mathbf{x}$ that better separates the classes is found, providing a decision surface in the form: $\mathbf{w}^T\mathbf{x}+c=0$, for a threshold c. Considering two normal multivariate distributed classes with means $\mu_\mathbf{1}$ and $\mu_\mathbf{2}$ and covariance matrices C_1 and C_2, respectively, the LDA approach consists in finding the weights $\mathbf{w}$ that maximize the ratio concerning the variance between the classes and the variance within the classes:

$$S=\frac{\sigma^2_{between}}{\sigma^2_{within}}=\frac{(\mathbf{w}^T(\mu_\mathbf{1}-\mu_\mathbf{2}))^2}{\mathbf{w}^T(C_1+C_2)\mathbf{w}}. \tag{2}$$

It is possible to show that this criteria is satisfied for $\mathbf{w}\propto(C_1+C_2)^{-1}(\mu_\mathbf{1}-\mu_\mathbf{2})$ and the threshold c is given by $\frac{1}{2}\mathbf{w}^T(\mu_\mathbf{1}+\mu_\mathbf{2})$. Once the training stage has been performed, trials with attribute vectors $\mathbf{x}$ are classified according to their position in the attribute space relative to the achieved decision hyperplane. The multi-class case was treated here analyzing all pairs of classes.

3 Results

Built in accordance with the experimental setup described in Sect. 2, Fig. 4 shows a typical DBINV map containing the main relevant features for class separation in spectral (Fig. 4a) and in recurrence-based attribute spaces (Fig. 4b), obtained for subject S1.

Some important conclusions can be drawn from the results shown in Fig. 4. For instance, it can be clearly noted that the electrode O9 was the one with best separation performance for both spectral and recurrence-based approaches for this specific

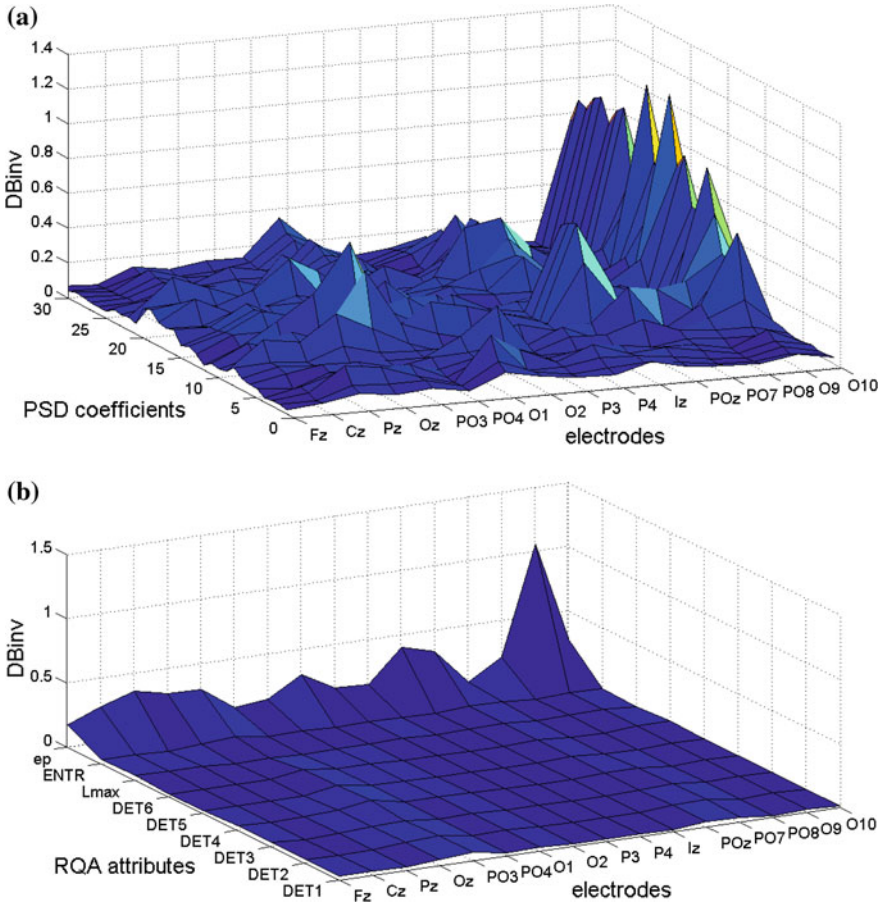

Fig. 4 Discriminant analysis using the inverse of the Davies-Bouldin index for **a** spectral and **b** recurrence attributes associated to S1. In this last figure, the threshold ε of the map is denoted by ep

subject. Effective separability was also attained with the IZ electrode in both attribute spaces. These locations completely agreed with the main cortical areas activated in SSVEP experiments, and, usually, the same promising electrode positions were identified in both attribute spaces analyzed here, despite there being a great variability concerning the achieved positions for different subjects, as commonly observed in BCI experiments.

Furthermore, Fig. 4b reveals that the threshold ε of the map was an effective attribute for separating the classes, the classical RQA measures having not reached a similar performance. This is a result observed for all subjects and probably follows from the adaptation of the recurrence plot in order to match a specific recurrence rate, a scenario that favors the adapted variable ε when comparing different recurrence patterns. The adoption of an adaptive recurrence plot is justified by the great

variability of the obtained signals, which made it difficult to find a general choice of a fixed ε for obtaining suitable recurrence plots.

Another particular RQA advantage was that the relevant features were much more concentrated—in terms of the number of attributes—in the recurrence scenario than in the spectral domain, which simplifies the automatic selection. For instance, if we consider class dispersion in the feature space, it is clear that the combination of the best PSD attributes with higher DBINV values could lead to correlated features with bad discrimination performance (Fig. 5a). On the other hand, the selection of the best PSD attributes considering different electrodes (selected by ranking the sum of all their respective PSD coefficients) can be more informative for separation (Fig. 5b). However, a clearly better separation scenario was attained in the recurrence-based attribute space as shown in Fig. 5c. In this case, note that both Fig. 5c and b used information from electrodes O9 and IZ, but class "down" (red points concentrated at the origin) was much more close to the "rest" state in the spectral domain, there also being intersections of other states in general. In fact, the classification error attained in Fig. 5b was 16.49 %, while in Fig. 5c this measure dropped to 3.81 %, i.e. 4 times

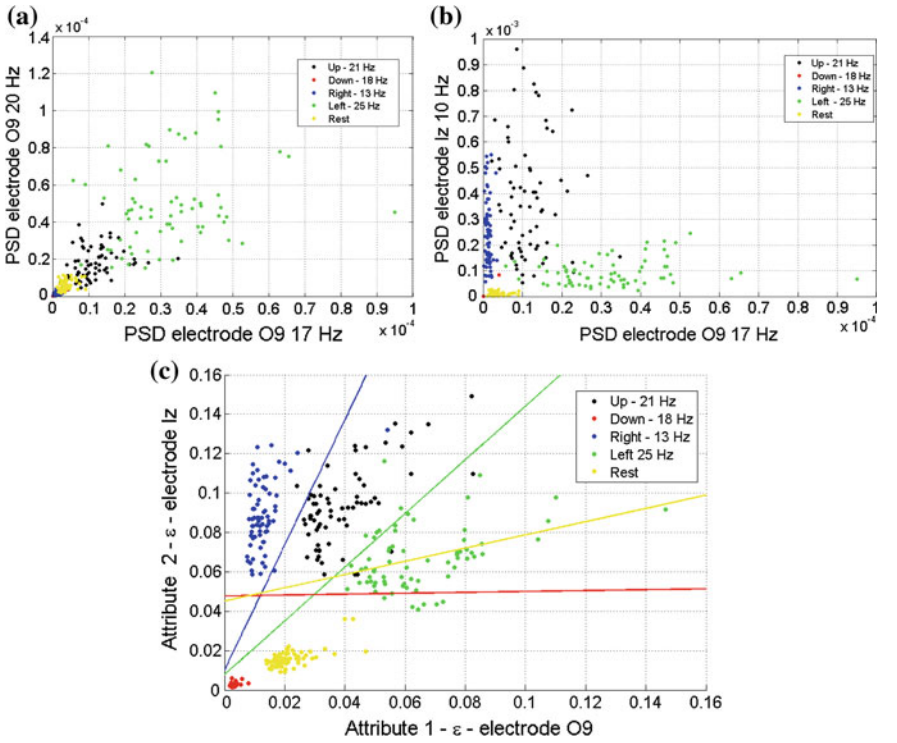

Fig. 5 Classes dispersion for **a** spectral attributes just ranked by DBINV measure; for **b** the best PSD attributes in the most promising electrodes according to DBINV value; and for **c** recurrence-based attributes domain associated to S1

lower. In this last case, the decision hyperplanes that separate class 1 ("up" state) from class i, with $i = 2, 3, 4, 5$ (in which the hyperplane color is associated with the color of the respective class), are shown, in order to illustrate the attribute space partition and that only few classification mistakes were made.

In addition to the classification performance attained here, the results described in Fig. 5 show that the recurrence analysis allowed the adoption of a simple heuristic for combining the features i.e. taking the attained ε for matching a specific recurrence rate ranked by DBINV index, while the analysis on the spectral domain would require a multi-objective optimization approach: a search that maximizes the DBINV measure, and, at the same time, minimizes the correlation between the selected attributes, which naturally poses a more complex task.

In order to investigate the classification performance in these different spaces, some simple heuristics for attribute selection were established. The first one consisted in simply taking the RQA features ranked by the DBINV index (black dots in Fig. 6). The second heuristic took the features ranked with the DBINV measure in the spectral domain (red dots in Fig. 6). The third ranked the electrodes according to the sum of the DBINV index of all frequency coefficients for that electrode and selected the PSD coefficients in the frequency of the visual stimulation (i.e. 13, 18, 21 and 25 Hz - blue dots in Fig. 6). The fourth heuristic ranked the electrodes such as in the third one and took the PSD coefficient with highest DBINV value for each electrode (green dots in Fig. 6). In all cases, the number of attributes was progressively increased and a k-fold cross-validation scheme was used to evaluate the mean classification error [13].

The evolution of the mean classification error can be observed in Fig. 6 for all subjects. It can be clearly noted that the recurrence-based attributes (specifically the threshold of the map that corresponds to the first 16 attributes) drastically dropped the classification error to lower than 10 % when 5 attributes were used and practically to 0 % with less than 8 features. In general, the performance achieved in the spectral domain was lower, with a mean classification error lower than 10 % attained only when more than 20 attributes were used in both second and third heuristics. The fourth heuristic took, progressively, the best PSD coefficients for different electrodes previously ranked, providing the better spectral scenario. Note that after using all the best coefficients of the 16 electrodes, this approach just attained a classification performance close to the recurrence scenario for S2 and a difference around 5 % for S1 and S2. In summary, using the best feature for different electrodes led to a faster convergence towards the minimal error and a better final performance in the recurrence-based attribute space.

Finally, a clear difference between the heuristics used for feature selection in the spectral domain can be also noted. Concerning heuristics 2 and 3, the free selection based on the DBINV score provided, at first, a better classification performance with a faster drop in the mean classification error, which was overcome by the selection based on the coefficients associated with the visual stimulation frequencies when few electrodes were taken into account (2 or 3 for all subjects). Interestingly, the best spectral scenario (heuristic 4) established that just the selection of the best PSD coefficient of each electrode could provide a low classification error (around 5 %) without requiring information concerning all the visual stimulation frequencies as

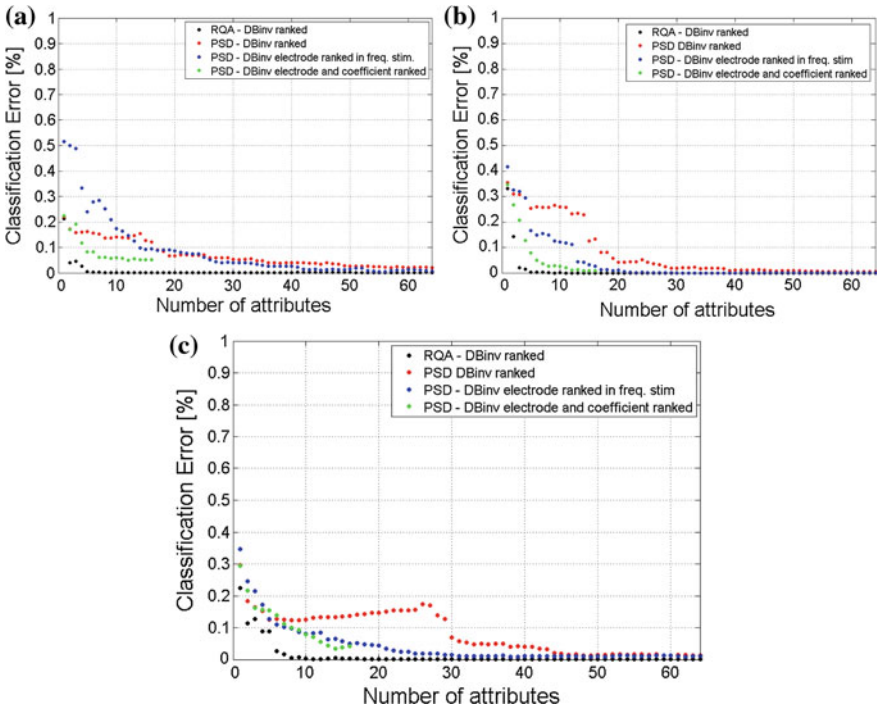

Fig. 6 Evolution of the classification error for a progressive increase in the number of attributes concerning the three aforementioned heuristics for subjects **a** 1, **b** 2 and **c** 3

commonly used in SSVEP- BCI systems. Such information could be helpful for designing BCI systems with just a few trials for training.

4 Discussion and Conclusions

In this work, the problem of efficient feature extraction for pattern recognition in SSVEP-based BCI was addressed. The obtained results clearly indicates the potential of using RQA in this task: it reached, indeed, a better performance than classical spectral analysis, commonly employed in this context. Curiously, it was found that the threshold of the recurrence plot was the key attribute to perform the class separation, which reveals the versatility of this complexity measure. In fact, if, on the one hand, the classical RQA measures are related to the diagonal structures in the RP and its generative deterministic rule [6], on the other hand, the recurrence rate is intrinsically associated with the quadratic Rényi entropy and other important information-theoretic measures can also be derived from it [6, 15].

In the present experiment, when the quadratic Rényi entropy is fixed (as a consequence of fixing the RR), the different attained recurrence plot thresholds reflect a particular metric of the recurrence structure of the signals that explain a fixed number of points. Note that this measure resembles that derived from the estimation of the correlation sum as introduced by [16] for the determination of the correlation dimension (D_2). In fact, the method used here just fixes the recurrence rate and makes use of the threshold for comparison between the different classes, instead of doing the opposite, when different thresholds define recurrence rates and a curve fit is used for obtaining D_2 in the classical approach [6, 16]. Interestingly, the procedure adopted here allowed overcoming the variability of EEG signals (and the recurrence plot drawbacks associated with it), and, at the same time, providing a better characterization of the temporal signal structure. The characterization of EEG signals in terms of the correlation dimension or nonlinear time series analysis has already been performed in the literature [17, 18], and the present work defines a particular application in the BCI context.

As a main drawback to the proposed recurrence-based approach, it can be mentioned the computational cost associated with the evaluation of an adaptive recurrence plot for this purpose, since BCI systems usually require a fast signal processing framework for real-time operation. Such limitation emphasizes the requirement for fast (dedicated) platforms for recurrence analysis, defining a natural extension to this work. The application of RQA for other BCI paradigms (e.g. motor imagery) also outlines a trend for future investigation.

Acknowledgments L.F.S. Uribe thanks CAPES for financial support. D.C. Soriano and R. Suyama thank UFABC and FAPESP (Grant number 2012/50799-2) for the financial support. F.I. Fazanaro thanks FAPESP (Grant number 2012/09624-4) for the financial support. R. Attux thanks CNPq for financial support. All authors are grateful to FINEP for funding the DesTine project (process number 01.10.0449.00).

References

1. Wolpaw, J., Wolpaw, E.W.: Brain-Computer Interfacing: Principles and Practice. Oxford University Press, New York (2012)
2. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002). doi:10.1016/S1388-2457(02)00057-3
3. Millán, J.R., Rupp, R., Mueller-Putz, G., Murray-Smith, R., Giugliemma, C., Tangermann, M., Vidaurre, C., Cincotti, F., Kubler, A., Leeb, R., Neuper, C., Mueller, K.R., Mattia, D.: Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. Front. Neurosci. **4**(161), 1–15 (2010). doi:10.3389/fnins.2010.00161
4. Sobani, Z.A., Quadri, S.A., Enam, S.A.: Stem cells for spinal cord regeneration: current status. Surg. Neurol. Int. **1**(1), 93 (2010). doi:10.4103/2152-7806.74240
5. Zhu, D., Bieger, J., Garcia Molina, G., Aarts, R.M.: A survey of stimulation methods used in SSVEP-based BCIs. In: Computational Intelligence and Neuroscience 2010, 702,357 (2010). doi:10.1155/2010/702357

6. Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. Phys. Rep. **438**(5–6), 237–329 (2007). doi:10.1016/j.physrep.2006.11.001
7. Marwan, N.: How to avoid potential pitfalls in recurrence plot based data analysis. Int. J. Bifurcat. Chaos **21**(4), 1003–1017 (2011). doi:10.1142/S0218127411029008
8. Zbilut, J.P., Zaldívar-Comenges, J.M., Strozzi, F.: Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data. Phys. Lett. A **297**(3–4), 173–181 (2002). doi:10.1016/S0375-9601(02)00436-X
9. Soriano, D.C., Suyama, R., Ando, R.A., Attux, R., Duarte, L.T.: Blind source separation in the context of deterministic signals. In: Eisencraft, M., Suyama, R., Attux, R. (eds.) Chaotic Signals in Digital Communications, pp. 325–348. CRC Press, Boca Raton (2013). doi:10.1201/b15473-13
10. Semmlow, J.: Signals and Systems for Bioengineers, 2nd edn. Academic Press, Amsterdam (2011)
11. Stoica, P., Moses, R.L.: Introduction to Spectral Analysis, Prentice-Hall, Upper Saddle River (1997)
12. Davies, D.L., Bouldin, D.W.: A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI–1**(2), 224–227 (1979). doi:10.1109/TPAMI.1979.4766909.
13. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 2nd edn. Academic Press, New York (1999)
14. Perez, J.L.M., Cruz, A.B.: Linear discriminant analysis on brain computer interface. In: IEEE International Symposium on Intelligent Signal Processing, 2007. WISP 2007, pp. 1–6 (2007). doi:10.1109/WISP.2007.4447590
15. Prichard, D., Theiler, J.: Generalized redundancies for time series analysis. Phys. D **84**(3–4), 476–493 (1995). doi:10.1016/0167-2789(95)00041-2
16. Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. Phys. D **9**(1–2), 189–208 (1983). doi:10.1016/0167-2789(83)90298-1
17. Acharya, R., Faust, O., Kannathal, N., Chua, T., Laxminarayan, S.: Non-linear analysis of EEG signals at various sleep stages. Comput. Methods Programs Biomed. **80**(1), 37–45 (2005). doi:10.1016/j.cmpb.2005.06.011
18. Stam, C.J.: Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin. Neurophysiol. **116**(10), 2266–2301 (2005). doi:10.1016/j.clinph.2005.06.011