

Ewa Piętka

Paweł Badura

Jacek Kawa

Wojciech Wieclawek *Editors*

Information Technologies in Medicine

5th International Conference, ITIB
2016 Kamień Śląski, Poland, June
20–22, 2016 Proceedings, Volume 2

Advances in Intelligent Systems and Computing

Volume 472

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within “Advances in Intelligent Systems and Computing” are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello, Universidad Central “Marta Abreu” de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at <http://www.springer.com/series/11156>

Ewa Piętka · Paweł Badura
Jacek Kawa · Wojciech Wieclawek
Editors

Information Technologies in Medicine

5th International Conference, ITIB 2016
Kamień Śląski, Poland, June 20–22, 2016
Proceedings, Volume 2

 Springer

Editors

Ewa Piętko
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

Jacek Kawa
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

Paweł Badura
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

Wojciech Wieclawek
Faculty of Biomedical Engineering
Silesian University of Technology
Gliwice
Poland

ISSN 2194-5357

ISSN 2194-5365 (electronic)

Advances in Intelligent Systems and Computing

ISBN 978-3-319-39903-4

ISBN 978-3-319-39904-1 (eBook)

DOI 10.1007/978-3-319-39904-1

Library of Congress Control Number: 2016940806

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

Preface

After many years of research and development in the area of computer-assisted diagnosis and therapy, multimodal medical imaging systems prove the ability to provide quantitative and qualitative information leading to increase the sensitivity and specificity of many clinical procedures. Often the aspect of multimodal analysis is extended to other domains including signals, video recordings, posture information, histopathological image quantification, patient tracking, and retrospective and prospective analyses, all combined with image data or one with another. Mathematical information analysis, computer applications together with medical equipment and instruments have become standard tools underpinning the current rapid progress with developing computational intelligence. However, modern systems still face challenges in performing hybrid data registration, subsequent analysis, and visualization to deliver pertinent quantitative information. We are witnessing a radical change as information technologies have been integrated into systems that address the core of medicine, including patient care in ambulatory and in-patient setting, disease prevention, health promotion, rehabilitation, and home care.

The *Fifth Conference on Information Technologies in Biomedicine* has become a recognized event that helps to bridge the gap between methodological achievements in engineering and clinical requirements in medical diagnosis, therapy, and rehabilitation. In these proceedings, members of academic societies of technical and medical background present their research results and clinical implementations. Many of these areas are recognized as research and development frontiers in employing new technologies in clinical environment.

Image and signal analysis are traditional parts that deal with the problem of data processing, recognition, and classification. Novel approaches to the 2D and 3D image analysis of various pathologically affected anatomical structures, early tumor detection, and stroke diagnosis are introduced. New research studies developed for the processing of heart, brain, and gastric as well as acoustic signals are discussed. Home care, remote rehabilitation assistance, and safety of the elderly require new areas to be explored in telemedicine, telegeriatics, and biomechatronics. Bioinformatics has become a dynamically developed field of computer-assisted

biological data analysis where information techniques are used in the nanoscale. Solutions of complex problems concerning diagnostics and therapy have been developed in the surgical procedures. A modeling and simulation platform presents also preimplementation studies required at the developmental stage of computer-assisted therapy.

We would like to express our gratitude to the authors who contributed their original research papers as well as to all reviews and *Scientific Committee Members* for their valuable comments.

Ewa Piętka

Scientific Committee Members

E. Piętka (Poland)
M. Akay (USA)
P. Augustyniak (Poland)
A. Bargieła (Great Britain)
R. Bruck (Germany)
K. Capova (Slovakia)
S. Czudek (Czech Republic)
A. Drygajło (Switzerland)
A. Gertych (USA)
D. Greenhalgh (Great Britain)
A. Grzanka (Poland)
M. Grzegorzek (Germany)
M. Gzik (Poland)
A. Hajdasiński (Netherlands)
J. Haueisen (Germany)
Z. Hippe (Poland)
M. Juszczuk (Germany)
E. Krupinski (USA)
M. Kurzyński (Poland)
A. Liebert (Poland)
A. Lipowicz (Poland)
R. Maniewski (Poland)
J. Marciniak (Poland)
M. McNitt-Gray (USA)
A. Mitas (Poland)
A. Napieralski (Poland)
E. Neri (Italy)
A. Nowakowski (Poland)
D. Paulus (Germany)
T. Pałko (Poland)
Z. Paszenda (Poland)

W. Pedrycz (Canada)
I. Provaznik (Czech Republik)
A. Przelaskowski (Poland)
P. Słomka (USA)
E. Supriyanto (Malaysia)
A. Świerniak (Poland)
R. Tadeusiewicz (Poland)
E. Tkacz (Poland)
H. Witte (Germany)
S. Wong (USA)
Z. Wróbel (Poland)

Contents

Part I Ambient Assisted Living

From Ambient Assisted Living to Active and Assisted Living: A Practical Perspective on Experiences and Approaches	3
Georg Aumayr	
Patients Assessment in Telegeriatrics Care	15
Jarosław Derejczyk, Jacek Kawa, Adam Bednorz, Joanna Lach, Aneta Hanusiak, Kinga Antoniak, Jan Szymshal and Piotr Seiffert	
Emotion Recognition Based on Physiological Sensor Data Using Codebook Approach	27
Kimiaki Shirahama and Marcin Grzegorzek	
Human Activity Recognition Using Smartphone Sensors	41
Marcin D. Bugdol, Andrzej W. Mitas, Marcin Grzegorzek, Robert Meyer and Christoph Wilhelm	
Assistive Integrated Personal Trainer System	49
Joanna Furmaniak, Roman Kaczorowski, Maciej Kasprzycki, Bartosz Zgrzeba, Zygmunt Kubiak and Ewa Łukasik	
Quantitative Validation of Gait and Swing Angles Determination from Inertial Signals	63
Paula Stepien, Zuzanna Miodonska, Agnieszka Nawrat-Szoltysik, Monika N. Bugdol, Michal Krecichwost, Pawel Badura, Piotr Zarychta and Marcin Rudzki	
Application of Gait Index Assessment to Monitor the Treatment Progress in Patients with Cerebral Palsy	75
Katarzyna Nowakowska, Robert Michnik, Katarzyna Jochymczyk-Woźniak, Jacek Jurkojć, Marek Mandra and Ilona Kopyta	

Recognising Gait Patterns of People in Risk of Falling with a Multi-layer Perceptron	87
Raoul Hoffmann, Christl Lauterbach, Axel Techmer, Jörg Conradt and Axel Steinhage	
Determination of Loads in the Joints of the Upper Limb During Activities of Daily Living	99
Agata Guzik-Kopyto, Robert Michnik, Piotr Wodarski and Iwona Chuchnowska	
Building Management System—Artificial Intelligence Elements in Ambient Living Driving and Ant Programming for Energy Saving—Alternative Approach	109
Jarosław Utracki	
Part II Bioinformatics	
Application of Adjoint Sensitivity Analysis to Parameter Estimation of Age-Structured Model of Cell Cycle.	123
Michał Jakubczak and Krzysztof Fujarewicz	
Multidimensional Correlated Mutation Analysis for Protein Contact Map Prediction	133
Kristyna Kupkova, Karel Sedlar and Ivo Provaznik	
Human Fibroblast Growth Factor 2 Hot Spot Analysis by Means of Time-Frequency Transforms	147
Anna Tamulewicz and Ewaryst Tkacz	
Structured Bi-clusters Algorithm for Classification of DNA Microarray Data.	161
Pawel Foszner and Andrzej Polanski	
Prediction of Newborn Weight Using Questionnaire Data and Machine Learning Approach	173
Justyna Pieter, Sebastian Student, Karolina Sobczyk and Krzysztof Fujarewicz	
Texture Analysis to Trophoblast and Villi Detection in Placenta Histological Images.	183
Zaneta Swiderska-Chadaj, Tomasz Markiewicz, Robert Koktysz and Wojciech Kozlowski	
Machine Learning Can Reliably Distinguish Histological Patterns of Micropapillary and Solid Lung Adenocarcinomas.	193
Nathan Ing, Sadri Salman, Zhaoxuan Ma, Ann Walts, Beatrice Knudsen and Arkadiusz Gertych	

Apoptosis in Cancer Cells	207
Eva Blahovcová, Henrieta Škovierová, Ján Strnádel, Dušan Mištuna and Erika Halašová	
Part III Modeling and Simulation	
Application of Reverse Engineering in Supporting the Treatment of Pectus Carinatum	217
Magdalena Antonowicz, Anita Kajzer and Wojciech Kajzer	
Breast Deformation Modeling Based on MRI Images, Preliminary Results	227
Marta Danch-Wierzchowska, Damian Borys and Andrzej Swierniak	
ToF-Data-Based Modelling of Skin Surface Deformation	235
Jan Juszczyk, Joanna Czajkowska, Bartłomiej Pycinski and Ewa Piętka	
Evaluation of Radiofrequency Field (2.4 GHz) Effects to Multilayer Structure of Human Skin and Pacemaker	245
Zuzana Psenakova, Maros Smondrk, Grazia Lo Sciuto and Mariana Benova	
Fluid Flow Analysis Using Finite Element Method, Determining the Effects of the Implantable Mechanical Heart Valves on Aortic Blood Flow	255
Przemyslaw Kurtyka, Witold Walke and Marcin Kaczmarek	
Forming an Occlusal Splint to Support the Therapy of Bruxism	267
Leszek Luchowski, Agnieszka Anna Tomaka, Krzysztof Skabek, Michał Tarnawski and Przemysław Kowalski	
Sensitivity Analysis of Signaling Pathways in the Frequency Domain	275
Małgorzata Kardynska and Jaroslaw Smieja	
Part IV Biomechatronics	
Development of New Testing Method of Mechanical Properties of Porcine Coronary Arteries	289
Bożena Gzik-Zroska, Kamil Jozzko, Wojciech Wolański and Marek Gzik	
Risk Factors Influencing Lower Limbs Injuries During IED Blast	299
Marek Gzik, Wojciech Wolański, Bożena Gzik-Zroska, Kamil Jozzko, Michał Burkacki and Sławomir Suchoń	

Hybrid Method of Human Limb Joints Positioning—Hand Movement Case Study	307
Grzegorz Glonek and Adam Wojciechowski	
A Survey of Selected Machine Learning Methods for the Segmentation of Raw Motion Capture Data into Functional Body Mesh	321
Magdalena Pawlyta and Przemysław Skurowski	
Evaluation of Automatic Calibration Method for Motion Tracking Using Magnetic and Inertial Sensors	337
Ewa Lach	
The Influence of Uncertainty in Body Segment Mass on Calculated Joint Moments and Muscle Forces	349
Magdalena Żuk and Celina Pezowicz	
Part V Biomaterials	
From the Research on Planning of Chemical Syntheses: Computer-Aided Modeling and Prediction of PGI2 Metabolic Reactions	363
Zdzisław S. Hippe	
Electrochemical Properties of TiO₂ Oxide Film on 316LVM Stainless Steel for Orthopedic Implants	375
Agnieszka Hyla and Witold Walke	
Evaluation of Physicochemical Properties of TiO₂ Layer on AISI 316LVM Steel Intended for Urology	385
Wojciech Kajzer, Anita Kajzer, Magdalena Grygiel-Pradelok, Anna Ziębowicz and Bogusław Ziębowicz	
Non-destructive Evaluation of Biomaterial's Magnetic Field After Plastic Deformation	399
Milan Smetana, Vladimír Chudacík, Klara Capova and Monika Oravcova	
Author Index	409

Part I

Ambient Assisted Living

The Special Session on *Adaptive Systems for Ambient Assisted* was organized as a part of the research projects: *Cognitive Village: Adaptively Learning Technical Support System for Elderly* funded by the German Federal Ministry of Education and Research, *My AHA: My Active and Healthy Ageing* funded by the European Commission (Horizon 2020), *Development of Innovative Informatics and Engineering Solutions in the Field of Medical Care* funded by the European Regional Development Fund under the Operational Programme Innovative Economy, and the *Telediagnostic Platform for Elderly People* founded by the National Centre for Research and Development (Poland).

From Ambient Assisted Living to Active and Assisted Living: A Practical Perspective on Experiences and Approaches

Georg Aumayr

Abstract AAL is a promising field of research for the last 10–15 years. As a topic, AAL has moved its focus from Ambient Assisted Living to Active and Assisted Living and is still moving to new directions. Under this perspective, the author summarizes experiences from projects like my-AHA (H2020) or AHEAD (AAL JP), SOFTCARE (AAL JP) etc. to provide a practical image and gives an overview of theoretical models in health care, data collection, measurements and methodological aspects on work with the elderly.

Keywords AAL · Frailty · Work with elderly · System theory · Salutogenesis

1 AAL Introduction and Connection to Health Care Providers

To start with the definition of Ambient Assisted Living, it is necessary to identify the related terms and changes of the last years. For example one of the major calls for this topic changed from Ambient Assisted Living to Active and Assisted Living. This has two implications: (a) AAL is a fluid term and (b) The focus changed from technological developments to user needs.

Especially the second implication has an interesting touch. In the beginning, AAL was a topic to develop new technologies and new products for a potentially emerging market. Developments were centred on the living space, fall detection and, more or less, about detecting emergencies. Already in a very early stage, first products reached the market (e.g. Personal Emergency Alert Systems), that are now established for almost 20 years. Safety was the major drive. But in the time between 2000 and 2010, no significant market breakthrough was achieved. New products were fancier but in their core had similar functions as in the 1990s. Since 2010 more and more new

G. Aumayr (✉)

Johanniter Austria Education and Research, Ignaz-Koeck Strasse 22, Vienna, Austria

e-mail: georg.aumayr@johanniter.at

URL: <http://www.johanniter.at>

products arrive at the market that are stuffed with new technology but are not reliable enough or produce too many false alarms. So they are not accepted very well by the end users. Another constraint was that developers believe that seniors could be afraid of new technology. This may have been true until the intuitive devices rolled over the market. Touchscreens provided a native way to interact with technology and the seniors were happy to find new ways to close up to their younger relatives in terms of communication technology. Another hit on the right spot were *exergames*. Exergames are the combination of exercises and gaming. With motion control of the Nintendo Wii, a new age of technology interaction was opened; even more, when Microsoft Kinect appeared on the market. This successes show why the term changed from Ambient Assisted Living to Active and Assisted Living.

Under the umbrella of AAL, developments for the elderly can be summarized. Including the use of wearables, new ways of video gaming, self-empowerment, communication etc., AAL became a synonym for technology having a focus on people with special needs. Those needs are as various as this field of development is. In times, where everything is under harsh cost control and the imperative of cost-efficiency, AAL is the bearer of hope to have an affordable health care in the future. AAL can support self-monitoring of patients and the gathering of health related data to be used to adjust therapy and to identify progress of decline or diseases. It also can help in improving quality of therapy when physicians and care givers are fully involved in the interpretation process. Health care is stated until now as the diagnostic of a status, a one shot observation. But this paradigm is changing. Health care is going to be a processual observation based on data. ICT-focused AAL is one way to collect this data and support and assist patients, physicians and all other care givers.

1.1 *Socio-Technical System*

A key to understand the importance of this field of action is the idea of a socio-technical perspective. Technology is not just the hardware and software but also the way it is integrated in culture and society. New technologies have to find a place in society. More detailed is this idea in the socio-technical theories by Weyer [24]. The two components of technic and society are linked to get the full meaning of a technology. Thinking about a similarity in social science, this is closely related to symbolic interaction [5] or the interactional view [23].

To define the needs of the target group, it is necessary to understand their motivation in testing and using technology. These are two fundamentally different things! In the AAL JP Project, a *motivational value approach* was developed, which is following an idea of economic motivation with the aim of self-evolvement. Value is thereby constructed by the individual in an interaction process with the social environment. Most interesting is, that this approach also takes into account that older people are volunteering for special tasks, only in the extent and the field of work, they wish to and, by this, to fulfil personal needs. The definition of needs as well as their hierarchy is according to the basic theory of needs by Maslow [14]. Following

this perspective, the first breakthrough on the market was related to the fulfilment of a basic need like personal security (e.g. MORE Project). Communication and social interaction is the next level of needs, which is explored and developed now to various extents, from video chat to social platforms. Mixed approaches are very promising in bridging the gaps between the levels of needs. One example would be the AAL JP Project AHEAD-Augmented Hearing Experience and Assistance for Daily Life. In this project hearing-glasses are combined with a health tracker and a smartphone with an app, which is providing health status information and compliance support with an integrated emergency function.

1.2 Recent Approaches

Recent approaches are following the idea to provide supportive information for self-management. By this, a lot of data is acquired and used for monitoring the health status and to identify tendencies and changes in behaviour to detect potential threats or changes of the health status to adjust therapy. The major aims of AAL nowadays are self-empowerment, personalizing healthcare and therapy and support the activities of the patient. By this, a general approach for activating the neighbourhoods of people in need of care and to integrate the informal caregiver became a supporting topic for AAL developments (e.g. AAL JP Project SOCIALCARE). An expression of this are the ongoing calls for proposals in H2020 Program. Topics addressed are chronic disease management and dementia and frailty prevention (e.g. H2020 Project my-AHA).

2 Data Collection and Data Use

Supporting with data became more of interest than giving physical support. One reason for this is the legal setting. Certain care related tasks are just allowed to nurses or physicians. So the support that can be given by technology is limited. But by collecting health related data, a more intensive support is possible for more than just one user group. Following the idea of a multi-angle development to address more target groups to increase acceptance, the secondary end users came into place and articulated their needs and provided expertise in care and health data monitoring.

2.1 hDATA

By this, it is necessary to take into account what data is collected for what purpose and how it is used in the end. To answer these questions for each project, it is helpful to use the hDATA approach [4].

Within this approach, a classification of data is available, which indicates the quality from raw data to interpreted data and shows the steps in between to use it in the context of health care:

- Basic data: heart rate, SpO₂, blood sugar, breathing frequency, etc.
- Processed data: heart rate variability, HbA1C, stress level etc.
- Interpreted data: diagnosis of disease.
- Class of understanding data: therapy design.

This processual perspective is also reflected in iterative cycles of therapy adjustments to the needs and developments of the patient. Aumayr [4] formalized in the hDATA cycle a draft for the understanding of this process according to the data classes. The hDATA cycle shows the diagnostic iterations in long term care from a practical perspective for inter-punctuation by the developer. It demonstrates the cybernetic (near to autopoiesis) understanding of health care as a process relying on items developed by its own system before hand and creates and adjust its upcoming steps (Fig. 1).

The circulation is focused on the principles of adherence and compliance, with the relationship between the patient and the health care professional (physicians and nurses) in the core. By this, a formalized set of relations and dependencies can be identified. The social environment is very important and should not be neglected. Every therapeutic action will be judged by the patient and his or her environment. To get a glimpse of the reach of the actions, a pragmatic overview of relations and their strength is shown in Fig. 2. The health system is centred on the relations between the patient and physician and consists of other relations and interdependencies. If a position is near to the centre, this means this relation has a high influence (strong relation) towards the patient-physician relation. If a relation is settled in the outer circles, the influence is weaker.

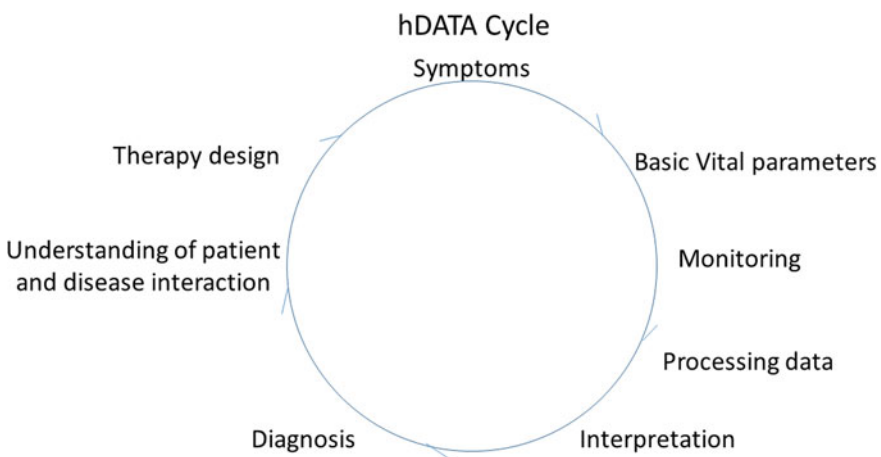


Fig. 1 hDATA cycle [4]

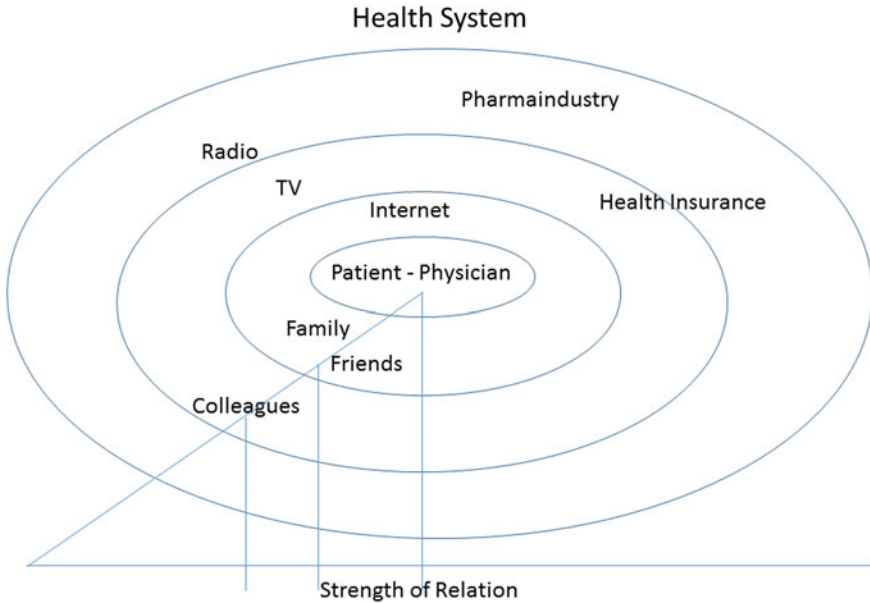


Fig. 2 Relations in health system [4]

To develop a new system, means also to know what kind of data is going to be collected, to which level it will be processed and to whom it will be transmitted. This has to be the basic setting for developments. A practical example is the measurement of activities of daily living. There are 6 activities, which can help to identify the level of support a person is in need of: Bathing, Dressing, Toileting, Transferring, Contenance and Feeding following the model by Noelker et al. [16]. These activities are trackable with a gyroscope and a location sensor. There are different acceleration patterns for each activity. But some aspects would be too similar to be identified by just the acceleration pattern. In combination with a localization sensor, to recognize in which room the person is, the patterns can be used to identify the activity. The most interesting thing is the detection of changes over a certain period in the available patterns. So it is possible to detect side effects of new medication, to detect if a person is losing capabilities to take care of themselves, or identifying new needs of support when the health status is decreasing. The needed sensors would be an accelerometer (3-axes) and a sensor for detecting the location indoors (e.g. NFC sensors at room entrances). The collected data would be available for the patient by a local interface, like a tablet or PC, and a report about the changes can be used to help the physician or the nurse in therapy adjustment.

In an even broader approach it would be possible to inform relatives about the activities of daily living or about changes in the daily routines to support them in their part of care. But the extent to which data is transmitted is always a discussion of ethics and law as well. As ethics are hard to define as settled, law is. By this, the

realization of the hDATA approach has to follow (1) the European Convention on Human Rights (2) the Council of Europe Convention 108 and (3) the European data protection law. The handling concerning data protection of all three of them and the principles to follow for data exchange and data use can be seen in the Handbook on European data protection law (www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf).

As long as the transparency is given and the primary end user, the patient, has signed and informed consent, a fair exchange between stakeholders is manageable and legal. A practical example is the design of the electronic health record ELGA in Austria with its health related telematics law.

3 System Theory Approach

In this context, health care is analysed as a system. There are several reasons for this perspective.

Health care is not just dealing with physical parameters but also with people and social aspects. These social aspects are founded on the relations the actors have to each other. By this, actions are interpreted from the individual point of view. The interpretation process provides information about the data as well as the contextuality. In the symbolic interactionism (1) people act to objects on the base of the meaning the object has to them; (2) this meaning is created by the social interaction with other people concerning the object; (3) this meaning is within the framework of constant discussion of the objects in an interpretation process and constantly used and adapted [5]. This makes the social interaction in health care a developing process. It is a construction and deconstruction of elements of use. By this, the interaction model becomes more and more complex and is building on its own experiences in the narrow environment. If a set of objects which build a closed cluster is defined, this can be understood as a system with self-referential behaviour [4]. Maturana and Varela describe such systems as autopoietic: *Autopoietic Systems are living organisms, which reproduce themselves. This happens as they build the components and elements, they are consisting of, from their own operations within their organization structure in an ongoing process. This can be imagined as components in a circular process with interacting iterations. By this the components are permanently produced to keep the system going* [15].

With the construction, deconstruction and reconstruction of relations and interpretations of system immanent elements, the overall system becomes more and more complex and stable. The dynamic element of change in the system is information. The existing structure in the health system is a top-down hierarchy from the physician to the patient. This disbalance of roles is accepted by the patient and is the foundation of the concept of compliance. By this, the full authority and responsibility is with the physician. In the last decades, this disbalance started to change. There are several reasons and factors for this. One of them is the easy access to information on the internet and the flood of publications in the health sector. By this, it is possible

that the disbalance between patient and physician is levelled by knowledge gain of the patient. Patients can become experts on their own health. When the hierarchy is broken and an equal understanding in the roles is given, the therapy is discussed between equals and the therapeutic decision is made by the physician and the patient as a team. The responsibilities are with the patient. This is called adherence.

In reality, most of the time responsibility is pushed towards the patient but the hierarchy remains in the mentioned terms of compliance. This produces (1) Frustration and (2) Reactance. The patients are not going to tell the physician about side effects of medication, just throw away the pills or change therapy themselves in an unsupervised way. This reduces the effectiveness of health care dramatically. It is interesting that a similar process was described in 1807 by Hegel in *The Phenomenology of the Mind* [10] for compliance and for adherence by Karl Marx in 1844 with the ideas of communism to break the concept of disbalance [13].

4 Salutogenesis

In this complex area it is not easy to keep the focus to just one thing. So many stakeholders and players are involved and have to be part of the development that the clear aim gets lost. To have a fundamental understanding of the principles of health care is a good support. As the term *health care* already proclaims, it is about taking care of the people's health, not their diseases or illnesses. This would be a therapeutically approach. In AAL the focus is on supporting people to live a life in autonomy and to stay in the environment they love and to support them to cope with changes in their capabilities. The theoretical approach of Salutogenesis by Aaron Antonovsky is acting in this way of thinking. The fundamental question of Salutogenesis [2] is: HOW DO PEOPLE STAY HEALTHY?

The basic idea follows the principles of a continuum between life and death. The movement to the pole of life would be health and to the pole of death would be sickness/illness. With a single measure it is not possible to identify the direction a person is tending to. But if several measurements and a processual view, it is possible to identify the direction and act before a point of no return is reached: This is called prevention. AAL is collecting a lot of data that can be used for tracking the health vector of the patient. By this, it is also possible to see effects of certain actions on the vector: This is called personalizing health and care. With this concept, it is possible to keep the aim in sight throughout a huge project. AAL can be a navigator to a healthier life.

5 Focus on Frailty

A movement to the pole of death is frailty. Frailty is a progressive process of increasing vulnerability, predisposing to functional decline, finally, leading to death. The loss of functional homeostasis, which is seen as the potential for the individual to

resist disease without loss of function, is one of the major criteria [22]. The effects of aging, lack of physical exercise, adequate nutrition, and inadequate defence to stressors lead to chronic under-nutrition, loss of bone substance and skeletal muscle mass. Frailty is operationally defined as a clinical condition meeting 3 out of 5 criteria [3] related to a physical phenotype including weak muscle strength, slow gait speed, unintentional weight loss, malnutrition or comorbidity, exhaustion and low physical activity. Some authors have added falls, delirium and urinary incontinence. Moreover, depression is a common feature of the aged, for which there is a specific geriatric depression scale, and which may in turn be both a consequence of frailty and a further cause of worsening. A scale of clinical frailty has been proposed, from 1 to 7, depending on the degree of dependence on others [21]. A pre-frail stage, in which one or two criteria are present, identifies a subset at high risk of progressing to frailty. Alternatively, frailty has been operationalized as a risk index by counting the number of deficits accumulated over time (termed *frailty index*, FI), including disability, diseases, physical and cognitive impairments, psychosocial risk factors, and geriatric syndromes (e.g. falls, delirium and urinary incontinence) [9]. Frailty is a construct of physical, neurological, psychological and social movements in the health continuum that are steering to the pole of death. Latest studies focused more on cognitive frailty and showed a high correlation between physical frailty and cognitive frailty. The proposed definition of cognitive frailty is a syndrome characterized by both physical frailty and cognitive impairment in absence of dementia. Cognition is now considered a relevant domain of frailty [19]. A recent study by Nouchi [17] showed that multi-domain intervention could improve or maintain cognitive functioning in elderly at-risk people of the general population. This is a good example for where assistive technology can help. Several sensors and questionnaires can be used to gather information and data on the individual patient as well as on a bigger population. With a frailty monitoring, early intervention can start and reduce long term effects to a minimum. It will not cure but it will decrease the speed of decline. This reduces the costs of a health system significantly.

6 How to Measure (GDS, TAM, CPQ-12)

To identify the potential of solutions, a set of measurements has to be taken. But to decide which scale is chosen, according to the impact that is aimed for, is up to the individual project. Nevertheless there are at least some scales that can improve the quality of research in AAL when used. Two major factors are interfering in research: Affinity to technology and personal traits like the Sense of Coherence. Cumulated with the recent health focus of the individual project, this already provides a long questionnaire. For the affinity towards and the experience with technology there are several validated tests available. Two examples are the Technology Acceptance Model (TAM) [8] and the Computer Proficiency Questionnaire (CPQ), which is especially designed for assessing seniors [6]. A problem of these questionnaires is their length. A shorter form is available with the CPQ-12. Another validated tool for

gather data about the experience with digital devices and services is the Technology Experience Questionnaire (TEQ) [7]. For measuring cognitive factors of frailty, the Geriatric Depression Scale (GDS) [26] is one possibility. There is even a shorter form available [12]. For the Sense of Coherence the long form is the Sense of Coherence Scale with 29 items (SOC-29). A shorter version is the SOC-13 [1], which is better suited to be part of a set of questionnaires.

This short overview of validated questionnaires that can be taken into account already shows a big potential of existing frameworks that can be used to check for certain variables. But there are some issues to be taken into account when working with seniors.

7 How to Measure with Seniors

Testing with real people instead of doing simulations is always a challenge. But with a very special, heterogenic target group, it is sometimes even more challenging, especially under the perspective of testing with vulnerable groups. The concept of vulnerability includes not just a physical aspect but also a social aspect and is a flowing state [25], similar to the idea of Salutogenesis. It is likely to test people twice and get very different results each time because they are in different periods of their life (e.g. change from working life to retirement, death of husband or wife etc.). Besides this component of dramatic changes in life, it also has to be taken into account that this target group has a very short lifespan. The dropout quote because of illness or death during long term studies is very high. When working with people over 80 years of age, a testing time of 3 months should not be exceeded [18]. It is also interesting, that people over 60 tend to refuse to take part in telephone-interviews more often than younger [11]. Resch and Aumayr summarized findings about testing with elder people and postulated an early approach for recruiting seniors as well as directly declare the benefit for them during the recruitment. This supports the commitment of the people and therefor reduces the effort for re-recruiting because of dropouts. By this, it is mandatory to use an understandable language, reduce the scientific terms to a minimum and explain them clearly [20].

8 Summary and Conclusions on AAL Research

AAL is a promising field for research. There are a lot of ways to address this topic in a decent way. But basic theoretical approaches have to be taken into account to keep the direction. An understanding of the health care sector and of psycho dynamics is recommended to find a decent perspective. But this also provides a huge area for research! Nevertheless, when discussing these issues, there are also limitations and lessons that we have learned to be taken into account. Major points are the recruiting and the planning with the needs of elder people to provide a good setting for trials

and interviews. Longitudinal studies can be done but a high effort in re-recruiting has to be expected. The target group is heterogeneous and not just describable by means of age. Activity levels and life style are more dominant when addressing a target group than demographics. Furthermore, the primary users are not the primary buyers. So the customers are mostly secondary end users, e.g. family members. For cost-intensive devices, the procurers are insurance groups or large care organizations, therefore business models have to address them.

But developments are paying off. In a time of demographic change to a society where the majority is of a higher age, costs in health care become higher and higher. The most cost intensive aspects are chronic diseases with medications and related stays in hospital. But medication is not as effective as it could be when it is a *standard medication*. A personalized medication is much more effective when adherence or at least compliance is in place. And hospital days can be decreased when a personalized therapy improves the effectiveness of interventions and decreases the number of long term effects and co-morbidities.

The request for low-cost devices and high price devices is given. The need for development is hidden in activities of daily life and the hierarchy of needs. The push by funding models is increasing, for research as well as for business. AAL is applied science from an early stage until hitting the market. It covers the development area to full length and is bringing social science, psychology, health care and engineering together, ICT being a bridge between the disciplines.

References

1. Antonovsky, A.: The structure and properties of the sense of coherence scale. *Soc. Sci. Med.* **36**(6), 725–733 (1993)
2. Antonovsky, A.: *Salutogenese. Zur Entmystifizierung der Gesundheit (Salutogenese. Unraveling the mystery of Health)*. dgvt Verlag, Tübingen (1997) (in German)
3. Arts, M., Collard, R., Comijs, H., Naude, P., Risselada, R., Naarding, P., Oude Voshaar, R.: Relationship between physical frailty and low-grade inflammation in late-life depression. *J. Am. Geriatr. Soc.* **63**(8), 1652–1657 (2015)
4. Aumayr, G.: Health data processing: system theory based meaning and potential for future health. *IDIMT-2015: Information Technology and Society Interaction and Interdependence*, pp. 423–430. Trauner Verlag, Pödebrady, Czech Republic (2015)
5. Blumer, H.: Der methodologische Standort des Symbolischen Interaktionismus (The methodological perspective of Symbolic Interactionism). In: A. B. (Hrsg.), *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit Bd.1*, pp. 80–146. Rowohlt, Reinbek (1973) (in German)
6. Boot, W., Charness, N., Czaja, S., Sharit, J., Rogers, W., Fisk, A., Nair, S.: Computer proficiency questionnaire: assessing low and high computer proficient seniors. *Gerontol.* **10**(9) (2013)
7. Czaja, S., Charness, N., Dijkstra, K., Fisk, A., Rogers, W., Sharit, J.: Computer and technology experience questionnaire (2006). http://create-center.gatech.edu/publications_db/report3ver1.3.pdf:CREATE
8. Davis, F.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **13**(3), 319–340 (1989)
9. Farhat, J., Velanovich, V., Falvo, A., Horst, H., Swartz, A., Patton, J.J., Rubinfeld, I.: Are the frail destined to fail? Frailty index as predictor of surgical morbidity and mortality in the elderly. *J. Trauma Acute Care Surg.* **72**(6), 1526–1530 (2012)

10. Hegel, G.: *Phaenomenologie des Geistes (Phenomenology of Mind)*. Suhrkamp, Frankfurt a.M. (1986) (in German)
11. Herzog, R., Kulka, R.: Interviewing older adults. A comparison of telephone and face-to-face modalities. *Public Opinion Q.* **47**, 405–417 (1983)
12. Leshner, E., Berryhill, J.: Geriatric depression scale-short form among inpatients. *J. Clin. Psychol.* **50**(2), 256–260 (1994)
13. Marx, K.: *Oekonomisch-philosophische Manuskripte. Kritik der Hegelschen Dialektik und Philosophie ueberhaupt (Economic-philosophical manuscripts. Critics of Hegel's dialectic and philosophy in general)*. (1844) (in German). <http://www.marxists.org>. Accessed 13 May 2015
14. Maslow, A.: *Motivation and Personality*. Harper and Bros, New York (1954)
15. Maturana, H., Valera, F.: *Autopoietische Systeme: eine Bestimmung der lebendigen Organisationen (Autopoietic Systems: a determination of living organisation)*. In: Maturana, H. (ed.) *Erkennen: Die Organisation und Verkoerperung von Wirklichkeit*, pp. 170–235. Vieweg, Braunschweig/Wiesbaden (1982) (in German)
16. Noelker, L., Browdie, R., Sidney Katz, M.D.: A new paradigm for chronic illness and long-term care. *Gerontol.* **54**(1), 13–20 (2014)
17. Nouchi, R.T.Y.: Four weeks combination exercise training improved executive functions, episodic memory and processing speed in healthy elderly people: evidence from a randomized controlled trial. *AGE* **36**(2), 787–799 (2014)
18. Panek, P., Höglinger, M., Hofer, K., Werner, K., Aumayr, G.: Anforderungen an ein tragbares Notrufgeraet fuer Senioren zur Foerderung eines sicheren Lebens zu Hause (Requirements on a wearable personal emergency device for seniors for improving safe living at home). In: Schreier, G., Hayn, D., Ammenwerth, E. (eds.) *eHealth 2011-Health informatics meets eHealth-von der Wissenschaft zur Anwendung und zurueck*, pp. 247–252. Wien (2011) (in German)
19. Qingwey, R., et al.: Cognitive frailty, a novel target for the prevention of elderly dependency. *Ageing Res. Rev.* **20**, 1–10 (2015)
20. Resch, K., Aumayr, G.: Methodische Herausforderungen bei Befragungen von und Testungen mit vulnerablen, aelteren Menschen ab 60 (Methodic challenges when interviewing of and testing with vulnerable, elder people 60+). In: Moser-Siegmeth, V., Aumayr, G. (eds.) *Alter und Technik*, pp. 129–142. facultas.wuv, Wien (2011) (in German)
21. Rockwood, K., Song, X., MacKnight, C., Bergman, H., Hogan, D., McDowell, I., Mitnitski, A.: A global clinical measure of fitness and frailty in elderly people. *CMAJ* **173**(5), 489–495 (2005)
22. Varadhan, R., Seplaki, C., Xue, Q., Bandeen-Roche, K., Fried, L.: Stimulus-response paradigm for characterizing the loss of resilience in homeostati regulation associated with frailty. *Mech. Ageing Dev.* **129**(11), 666–670 (2008)
23. Watzlawick, P., Beavin, J., Jackson, D.: *Menschliche Kommunikation. Störungen. Paradoxien (Human communication. Disfunction. Paradoxes)*. Huber, Bern (1969) (in German)
24. Weyer, J.: *Techniksoziologie. Genese, Gestaltung und Steuerung sozio-technischer Systeme (Technicsociology. Genesis, Design, and Steering of socio-technical Systems)*. Juventa (2008) (in German)
25. Witteck, C.: *Die Bedeutung der Resilienzforschung fuer die Sozialpaedagogik (The meaning of resilience research for socio-paedagogic)*. Grin Verlag fuer akademische Schriften, Norderstedt (2008) (in German)
26. Yesavage, J., Brink, T., Rose, T.L. et al.: Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiatr. Res.* **17**(1), 37–49 (1982)

Patients Assessment in Telegeriatrics Care

Jarosław Derejczyk, Jacek Kawa, Adam Bednorz, Joanna Lach,
Aneta Hanusiak, Kinga Antoniak, Jan Szmyszal and Piotr Seiffert

Abstract Elder patients are considered to have a high degree of frailty and active multiple pathology. When access to geriatric care is very limited, diagnostic methods without face to face contact are of prime importance, especially during initial patient assessment. The goal of the paper is to present and validate telemedicine screening procedures adopted in the John Paul II Geriatric Hospital (GH), covering the neurological part of Comprehensive Geriatric Assessment (CGA). The participants of the study have been patients aged 65 and older. Three methods have been evaluated: Telephone Interview for Cognitive Assessment (TICS), Geriatric Depression Scale (GDS) questionnaire filled out during video call, and analysis of a written text. The third procedure aims at detecting cognitive skills, providing alternative to Montreal Cognitive Assessment (MoCA) and Mini Mental State Examination (MMSE). Face-to-face methods are used as a gold standard reference. High or moderate correlations have been observed in all cases, i.e. in TICS versus reference MMSE scores (Spearman's $R = 0.7955$), in face-to-face GDS versus videoconference Tele-GDS (Spearman's $R = 0.7148$), and in the written text versus MoCA. (Spearman's $R = -0.6649$). The results of this study presume that TICS can be applied as a reliable method of remote diagnostics. In the two latter cases, the use of telegeriatrics has not been excluded, and further research is planned. Comprehensive assessment has been considered to be crucial for delivering geriatric care. Work in the field of telemedicine-based comprehensive assessment of geriatric patient should be continued.

Keywords Telegeriatrics · Telemedicine · Comprehensive geriatric assessment · MoCA · MMSE · GDS · Analysis of writing

J. Derejczyk · A. Bednorz · J. Lach · A. Hanusiak · P. Seiffert
John Paul II Geriatric Hospital, Katowice, Poland

J. Kawa (✉) · K. Antoniak
Faculty of Biomedical Engineering,
Silesian University of Technology, Roosevelta 40, Zabrze, Poland
e-mail: jacek.kawa@polsl.pl

J. Szmyszal
Faculty of Materials Engineering and Metallurgy, Silesian University of Technology,
Krasińskiego 8, Katowice, Poland

1 Introduction

Access to experienced physicians, psychologists and psychotherapists specialized in gerontology is becoming increasingly necessary for both medical and non-medical facilities aiming at provision of effective diagnosis and treatment of patients at advanced age. Yet, in practice, the access to geriatric care in some countries is very limited, for example there were only 333 specialists of geriatrics in Poland in 2014 [31] (i.e. 0.9 specialist per 100,000 citizens). The ratio of beds per 100,000 citizens in Poland is also low: 1.8 [22], especially when compared to neighboring Slovakia (14 beds [10]) or Belgium (with the highest ratio of over 100 beds per 100,000 citizens [9]). It is expected that fast access to powerful diagnostic procedures could nevertheless be secured with the use of telegeriatric-based solutions. In recent years, various initiatives have been started in the field of telemedicine in geriatric care (telegeriatrics). Levine et al. [19] shows that smartphone applications can assist in chronic, non-cancer pain treatment. Lichstein et al. [20] uses Skype to help insomniac patients. Rajasekaran et al. [25] employs remote monitoring of patients during drug therapy. Many projects are dedicated to helping diabetics and people with vascular diseases [32]. Mixed, traditional and telemedicine-based, services are provided for selected patients [3]. However, the available solutions are mostly dedicated to patients already diagnosed, and lack comprehensive approach needed during initial contact with various gerontological specialists.

The geriatric diagnostic process begins with a thorough history taking and physical examination, which are aimed at separating the natural signs of aging from existing diseases and defining the person's functional capacity. Diagnosis in geriatrics is difficult because one patient may suffer from many overlapping diseases [7]. This may lead to severe and immediate disabilities, such as stroke, as well as progressive disabilities that slowly erode the ability of the elderly to care for themselves.

When it comes to geriatric diagnostics, besides laboratory testing and imaging, different scales and tests are being used, such as Barthel index, ADL scale (activity of daily living) [8] or Vulnerable Elder Survey [29]. Selected patients are subjected to further functional and physiological evaluation with the use of complex and interdisciplinary tools such as comprehensive geriatric assessment (CGA) [30].

The CGA includes a number of screening tests as well as evaluation of daily living activities, social support, cognition, mood, gait and falls, nutrition, sensory impairments, incontinence, polypharmacy, biochemical results, abuse of the elders, pressure ulcers, pain and advanced directives [6].

Our recent projects have been focused on developing a telemedical counterpart of comprehensive geriatric assessment (CGA), comparable with the traditional approach. Expected benefits include increased access to geriatric care and extended coverage of screening procedures.

2 Materials and Methods

Telemedicine procedures included in the paper are intended as a remote counterpart of the neurological part of CGA. The primary use scenario involves geriatric specialists consulting patients who stay at medical or non-medical facility. The subjects remaining under care of trained staff are examined by means of telecommunication devices by a specialist who controls the procedure.

For the purpose of the study, during evaluation of each of the three verified procedures, first the group of possible candidates was selected from John Paul II Geriatric Hospital in Katowice (GH) patients, on the basis of medical documentation. The groups of patients selected for each procedure might have overlapped, but they were not the same. The patients (or, in specific cases, legal guardians) were asked to participate, and their initial consent was obtained. In the next step, the staff of GH was instructed to talk with the patients, outline the details of the procedure and provide necessary information.

The proposed set of procedures evaluated in GH has been chosen on the basis of relevance versus the traditional face-to-face neuropsychological part of CGA. The methods do not cover the whole examination, yet provide powerful screening capabilities.

All the presented work has been approved by the Bioethics Commission of the Academy of Physical Education in Katowice, Poland (resolution 1/2010, with amendments).

The first procedure tested has been the Polish version of TICS (Telephone Interview for Cognitive Assessment). In its original version, the observed screening abilities have been comparable [26, 28] with the face-to-face Montreal Cognitive Assessment (MoCA) [11, 12] and Mini-Mental State Examination (MMSE) [23] examinations. However, this had not been confirmed before for all the local versions of TICS. In the study, a Polish version of the TICS has been evaluated. The Polish version of TICS evaluated in GH follows the version of the tests presented in English [2, 16, 33] in both the type of questions and the grading. The cut-off level was established at 30 points. The examination was executed by performing three tests in one diagnostic session with a neuropsychologist. The gold standard MMSE and MoCA tests were performed in the usual manner—face to face in typical medical office conditions. The results of this evaluation have partially been presented in our conference paper [5].

The second procedure has been Tele-GDS (Telemedicine Geriatric Depression Scale). Remote methods of diagnosis, which is easy for patients, are considered important worldwide (cf. Gellis et al. [13]). In their teliagnostic form, Hamilton Depression Rating Scale (HAM-D) [14] and Brief Psychiatric Rating Scale (BPRS) [24] were among the first methods to be evaluated [17] by comparison of results of face to face and video consultation/examination. In 2001, Menon [21] presented a similar evaluation of Geriatric Depression Scale (GDS) [35] and Hamilton Depression Rating Scale, over a low quality video link. In this paper, the procedure has been evaluated using high quality video. GDS, in the form proposed by Yesavage

[34, 35], has been evaluated in GH using video equipment. The used variant of GDS has been graded in the range of 0–15 pts., where depression with increasing intensity has been diagnosed in case of patients with six or more points scored [35]. All the patients were informed of the purpose, the course of the study and security policy regarding the data acquired. After giving his/her consent, the patient participated in the study conducted in typical medical office conditions, with the first psychologist performing MMSE and MoCA tests (to assess cognitive functions), and standard Geriatric Depression Scale rating. In the next phase, a day or more later, a second psychologist performed Tele-GDS (the same GDS rating using video conference equipment). Additionally, in several cases, the procedures were performed in the opposite order (Tele-GDS procedure first).

The third procedure, experimental analysis of writing, aims at detecting cognitive skills providing simple and fast screening alternative to MoCA and MMSE, and replacing TICS in selected scenarios of remote diagnosis. Typical examination in MMSE takes about 20 min (no impairment or mild cognitive impairment), while MoCA is longer, with examination time of about 25 min. A typical session of TICS assessment takes 20 min. In all cases, constant supervision of the examiner is necessary to provide instructions and grade answers. During diagnostic sessions in GH, a hypothesis has been formed that a less time and resource consuming version of cognitive screening test could be proposed, which would be based solely on subject's writing and could be performed both locally or using telecommunication devices. This corresponds—to some extent—to other research, conducted on writing and mental condition, e.g. handwriting versus depression [27], handwriting in Parkinson's disease [4] etc. In this paper the procedure used in a pilot study and its evaluation has been included.

The subjects were informed about the goal of the study and they gave consent. First, the MoCA and MMSE tests were performed face to face. Next, Geriatric Depression Scale Assessment (not included in this evaluation) was performed. Finally, the subjects were asked to perform/complete two exercises with direct, face to face assistance of the examiner. In each case the examiner addressed the subject directly (in Polish) and waited for the completion of the task:

1. Three sentences will be read aloud and then dictated to you. Write them down on a blank piece of paper. These are the sentences: I am at hospital. I have received a blank piece of paper and then a pen. I am writing the text down on the paper received.
2. I am going to read to you twice a single sentence. Please write it down on the received piece of paper using only capital letters. This is the sentence: I wonder what the weather is going to be like tomorrow. In several cases, a third exercise was also included:
3. Draw two, similar-size circles: one over another. Mark the center of both of the circles.

During the first phase of the study, ordinary pen and white paper were used. Later, both regular and digital writing registration were employed using Livescribe Echo Pen that replaced the ordinary pen of similar size, used earlier. Such registered data

were encoded using libsmartpen library and analyzed. Using the collected samples of writing, a series of writing parameters have been extracted manually and with the use of customized Matlab software. The examinations were performed in consulting room environment, with natural and artificial light.

3 Results

Three methods: TICS, GDS and analysis of writing were evaluated using statistical tools and appropriate reference methods. The results of the evaluation are presented below.

3.1 Telephone Interview for Cognitive Assessment

A group of 55 people, patients of GH, was tested. The study group consisted of 55 patients, the average age was 81.2 years, 42 patients were women (average age 81.1 years) and 13 were men (average age 81.5 years) [5]. All the tested patients agreed for the extended procedure and were interviewed on the telephone. All of them were over 75 years old, did not take drugs with pro-cognitive effects, were not diagnosed with dementia, and—among others—were able to participate in MMSE and MoCA test used here as gold standard. Various environmental and health data have been gathered. The results have been verified statistically. No correlation has been found (Spearman rank analysis) between TICS and the age of patients, education history, diabetes, depression or Parkinson's disease, etc.

Pearson correlation and Spearman rank correlation between the results of MMSE and MoCA tests were $R = 0.8085$ ($p < 0.0001$), $R_{SPEARMAN} = 0.8170$ ($p < 0.0001$), and the correlation between the results of MMSE and TICS tests was $R = 0.7867$ ($p < 0.0001$), $R_{SPEARMAN} = 0.7955$ ($p < 0.0001$) between MoCA and TICS tests: $R = 0.7741$ ($p < 0.0001$), $R_{SPEARMAN} = 0.7684$ ($p < 0.0001$). Figure 1; i.e. high correlation has been observed between the gold standard MMSE and MoCA and TICS tests.

As pointed out by the subjects, fatigue was noticeable during TICS part, as all the tests: MMSE, MoCA and TICS were performed in one go.

3.2 Geriatric Depression Scale Assessment (Tele-GDS)

25 patients were examined during the evaluation. The first group of 17 patients participated in the study in 2014 (partial results have been discussed in our previous paper [1]), while the second group of 8 patients was additionally examined in 2015.

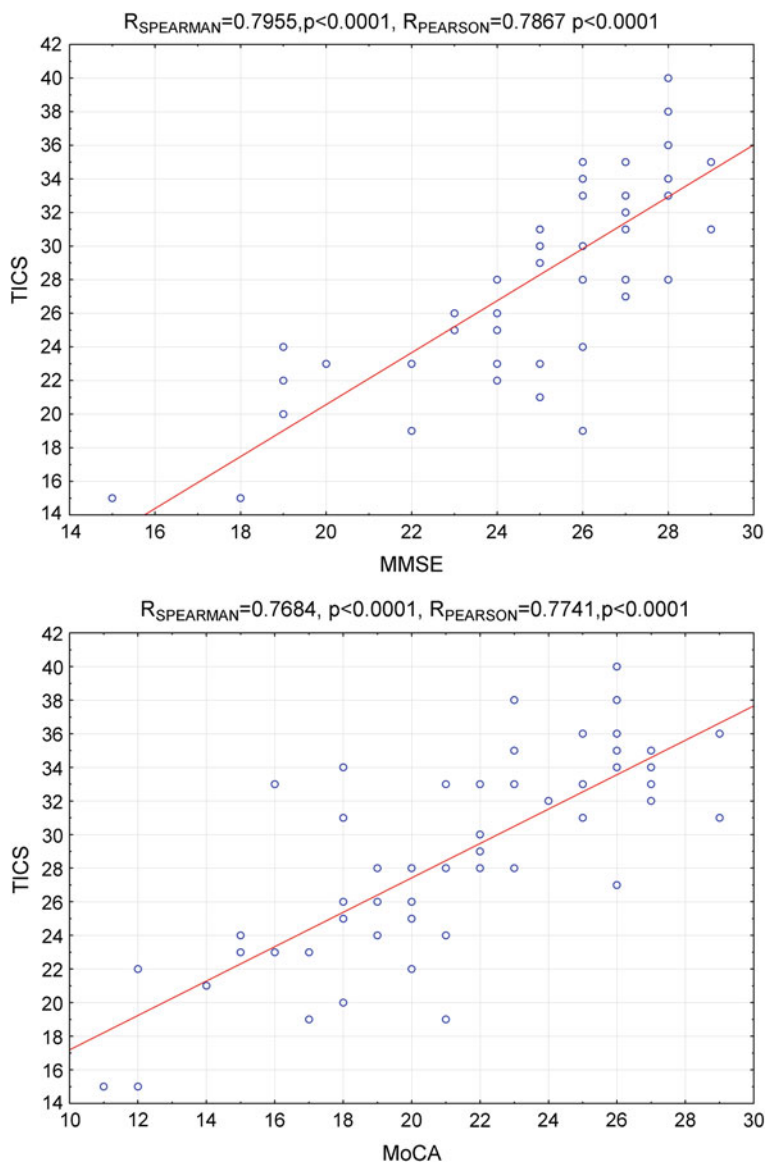


Fig. 1 Correlation between results of MMSE versus TICS test (*top*) and MoCA versus TICS test (*bottom*)

In 21 cases, the traditional face to face examination preceded the Tele-GDS, while in 4 cases the order was reversed. Statistical evaluation has been performed.

The mean GDS result for reference face-to-face assessment was 4.3 pts (SD 3.3 pts) and for videoconsultation Tele-GDS—3.5 (SD 2.6 pts.) Pearson correlation

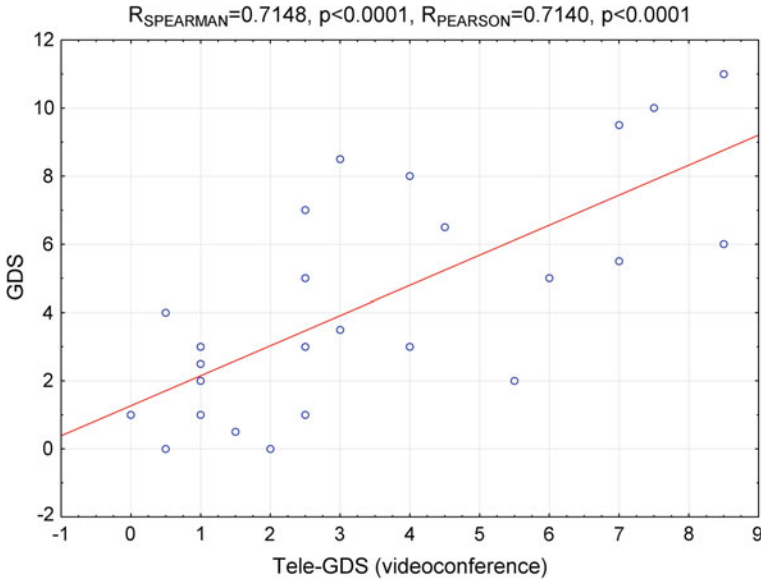


Fig. 2 Linear correlation for GDS and videoconsultation GDS

between both GDS assessments was $R = 0.7140$ ($p < 0.0001$) (Fig. 2). Spearman rank correlation between both GDS assessments has been calculated as $R_{SPEARMAN} = 0.7148$ ($p < 0.0001$). Despite the correlation, in four cases patients diagnosed with depression in face to face examination have not received the same diagnosis by videoconsultation (type II error). As shown in Fig. 2, the results obtained were lower mostly in videoconsultation. Assuming the same threshold for both tests, established for traditional GDS [34] (six or more points), sensitivity and specificity have been assessed as 56 % and 88 %, respectively.

Various factors have been considered, that may lower the results of GDS in videoconsultation, including: (1) insufficient perception of questions due to the use of telecommunication equipment (2) lack of direct contact, weakening the relationship between the examiner and the patient, (3) second examination being performed too soon after the first one, and the subject experiencing the fear of evaluation, which resulted in her/his dissimulation. However, the lower results during Tele-GDS have also been observed during sessions, in which Tele-GDS was the first test (2 of 4 cases). Further investigation is necessary to explain this phenomenon.

In the interview following the examination, several advantages of videoconsultation have been noted by psychologists: (1) it is easier to establish contact with the patient, (2) the examination is easily accessible, (3) facial expression can be observed in great detail, also in recordings, (4) patients, who may not be able to participate in face to face examination, due to physical limitations, may be examined, and (5) examinations can be repeated and results easily compared using recordings. The disadvantages pointed out include: (1) the fact that unnoticed hearing impairment

may affect the results, (2) the behavior of patients cannot be observed as well as during direct meetings; some non-verbal communication may be missing, (3) patients may be overwhelmed by technology and strongly prefer traditional examination; (4) additional consent is necessary.

3.3 Analysis of Writing

The pilot study [18] was conducted between February and May 2014. 26 patients of GH (avg. age 76 years, 20 women, 6 men) were examined. Making use of the obtained samples of writing, several features have been extracted. The presence of tremor, level of focus, number of pen detachments (exercises with written and block letters separately), etc. have been assessed for all samples. Time of writing, angle and various dynamic parameters have also been extracted for six electronically registered samples. The number of pen detachments has been chosen for correlation with MoCA and MMSE. One case was excluded from MoCA analysis, as the test was not completed during the session.

The negative correlation between the number of pen detachments in written text and MMSE was $R = -0.6146$ ($p = 0.001$), while for written text and MoCA it amounted to $R = -0.5509$ ($p = 0.004$), Fig. 3. The calculated Spearman rank correlation for the number of pen detachments in written text and MoCA was $R_{SPEARMAN} = -0.6649$, ($p < 0.001$), and in case of written text and MMSE $R_{SPEARMAN} = -0.6167$ ($p < 0.001$). No correlation has been confirmed to exist between pen detachments in block letters and gold standard.

The obtained mild correlation values between the number of pen detachments in reference text and the results of MoCA and MMSE are considered promising, but the evaluated number of samples should be increased to confirm the screening value.

3.4 Discussion

High correlation between the Polish version of TICS and MoCA/MMSE tests confirms the screening value of the local, Polish version of TICS. Mild correlation of Tele-GDS performed during videoconsultation and regular GDS also confirms the screening value, however, further investigations are needed, following the noticeably lower average results obtained during the videoconsultation test. Experimental analysis of writing samples is also promising, but further work is necessary in order to collect and analyze more data.

In our opinion, work in the field of comprehensive assessment of geriatric patients with the use of telecommunication should be continued. Implementing telemedicine solutions in mental health examination improves diagnostics and access to care in most scenarios [15], and despite age-related digital divide, the effects in geriatrics should be similar.

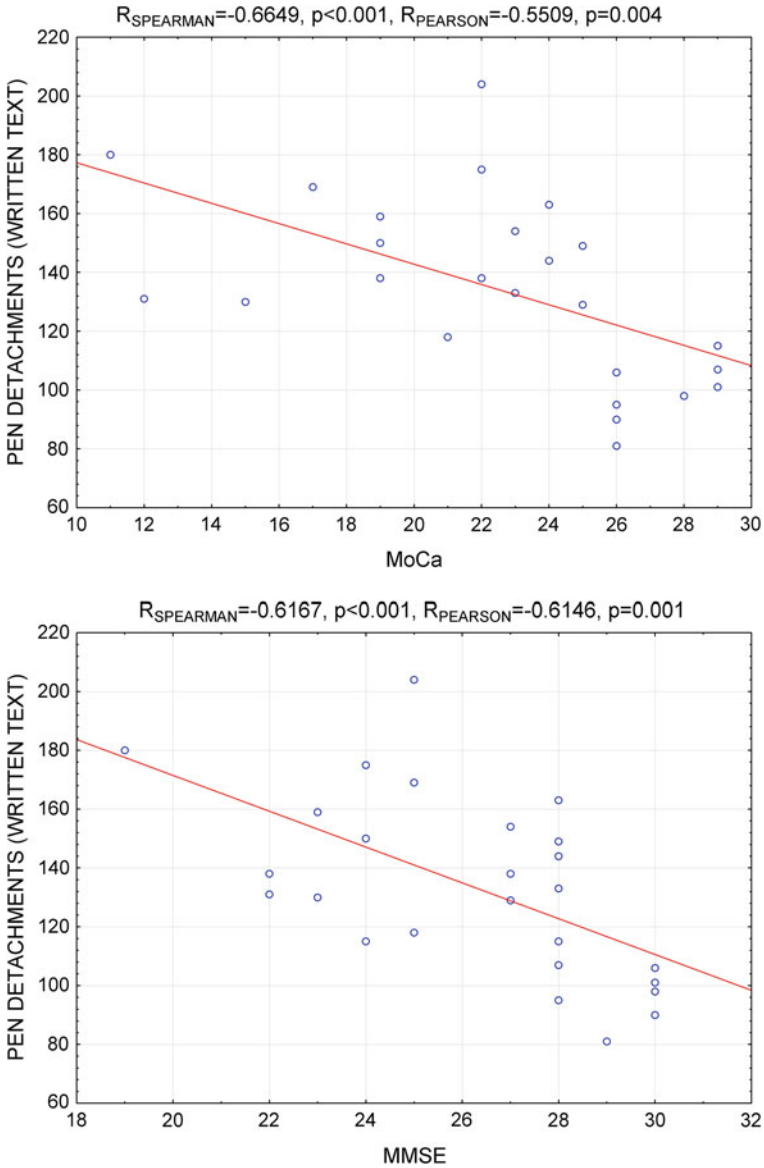


Fig. 3 Correlation between number of pen detachments in written text and MoCA (*top*) and MMSE (*bottom*)

It is crucial in Poland, where despite the aging society and growing demands, the number of geriatric specialists and hospital beds is very low. The use of telegeriatrics may be the only option available, especially in rural areas. However, remote access

to fast screening procedures should provide several benefits, even in countries with more developed elder care:

- residents of nursing homes and similar non-medical facilities can be easily examined, which increases the coverage of screening procedures,
- elder inhabitants of isolated locations and communities can be included in screening programs,
- access to geriatric care in non-geriatric medical facilities may be secured.

Once the vulnerable patient is identified using telegeriatrics, s/he can later be examined and treated using conventional means.

Consequently, further research has been planned in GH, especially in Tele-GDS and writing samples analysis.

Other popular geriatric tools that could be used in remote manner (e.g. MiniCOG) will be included.

Acknowledgments The research has been partially co-financed by the European Regional Development Fund under the Operational Programme Innovative Economy (project no. POIG.01.03.01-24-061/12) as well as National Centre for Research and Development (project no. IS-2/54/NCBR/2015).

References

1. Bednorz, A., Derejczyk, J., et al.: Zastosowanie wideokonsultacji w rozpoznawaniu depresji u osób starszych, jako próba wykorzystania telediagnostyki w geriatricii (Use of videoconsultation in depression diagnostics in elderly people as an attempt of application of telediagnosics in geriatrics). *Geriatrics* **8**, 141–149 (2014). (in Polish)
2. Brandt, J., Spencer, M., Folstein, M.: The telephone interview for cognitive status. *Cogn. Behav. Neurol.* **1**(2), 111–118 (1988)
3. Cardozo, L., Steinert, J.: Telemedicine for recently discharged older patients. *Telemed. J. E. Health* **16**, 49–55 (2010)
4. Cobbah, W., Fairhurst, M.: Computer analysis of handwriting dynamics during dopaminergic tests in Parkinson's disease. In: *Euromicro'00*. pp. 2414–2418 (2000)
5. Derejczyk, J., Stępień-Wyrobiec, O., Hanusiak, A., et al.: The application of adapted TICS (Telephone Interview for Cognitive Status) for diagnostics of cognitive function disturbances in elderly patients. a pilot study. In: Pietka, E. et al. (eds.) *Information Technologies in Biomedicine*, vol. 4, pp. 115–123. Springer (2014)
6. Devons, C.A.: Comprehensive geriatric assessment: making the most of the aging years. *Curr. Opin. Clin. Nutr. Metab. Care* **5**(1), 19–24 (1 2002)
7. Dimitrova, R.: Growth in the intersection of eHealth and active and healthy ageing. *Technol. Health Care* **21**, 169–172 (2013)
8. Elsayy, B., Higgins, K.: The geriatric assessment. *Am. Fam. Phys.* **1**, 48–56 (2011)
9. European Union Geriatric Medicine Society: Belgische vereniging voor gerontologie and geriatricie (2008). www.eugms.org/our-members/national-societies/belgium.html. Accessed Nov 2015
10. European Union Geriatric Medicine Society: Slovak gerontology and geriatric medicine society (2005). www.eugms.org/our-members/national-societies/slovakia.html. Accessed Nov 2015
11. Folstein, M., Folstein, S., McHugh, P.: Mini-mental state. *J. Psychiatr. Res.* **12**(3), 189–198 (1975)

12. Folstein, M., Robins, L., Helzer, J.: The mini-mental state examination. *Arch. Gen. Psychiatry* **40**(7), 812 (1983)
13. Gellis, Z., Kenaley, B., Have, T.: Integrated telehealth care for chronic illness and depression in geriatric home care patients: the integrated telehealth education and activation of mood (I-TEAM) study. *J. Am. Geriatr. Soc.* **62**(5), 889–895 (2014)
14. Hamilton, M.: A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* **23**, 56–62 (1960)
15. Hilty, D.M., Ferrer, D.C., Parish, M.B., Johnston, B., Callahan, E.J., Yellowlees, P.M.: The effectiveness of telemental health: a 2013 review. *Telemed. J. E. Health* **19**(6), 444–454 (2013)
16. Hogervorst, E., Bandelow, S., Hart, J., Henderson, V.W.: Telephone word-list recall tested in the rural aging and memory study: two parallel versions for the TICS-M. *Int. J. Geriatr. Psychiatry* **19**(9), 875–880 (2004)
17. Jones, B., Exum, M., McFarlane, M.: Telemedicine ratings of geriatric depression. *Int. J. Psychiatry Med.* **26**, 392 (1996)
18. Kužel, K.: Analiza pisma w diagnostyce osób starszych (Handwriting analysis in diagnosis of elder people). Master's thesis, Silesian University of Technology, Poland (2014) (in Polish)
19. Levine, M., Richardson, J.E., et al.: Novel telemedicine technologies in geriatric chronic non-cancer pain: primary care providers' perspectives. *Pain Med.* **15**(2), 206–213 (2014)
20. Lichstein, K.L., Scogin, F., Thomas, S.J., et al.: Telehealth cognitive behavior therapy for co-occurring insomnia and depression symptoms in older adults. *J. Clin. Psychol.* **69**(10), 1056–1065 (2013)
21. Menon, A., Srikumar, M., et al.: Evaluation of a portable low cost videophone system in the assessment of depressive symptoms and cognitive function in elderly medically ill veterans. *J. Nerv. Ment. Dis.* **189**, 399–401 (2001)
22. Nałecz, S. (ed.): Health and Health Care in 2012: Statistical Information and Elaborations (2013). https://old.stat.gov.pl/cps/rde/xbcr/gus/HSW_health_and_health_care_2012.pdf
23. Nasreddine, Z.S., Phillips, N.A., Bédirian, V., et al.: The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**(4), 695–699 (2005)
24. Overall, J.E., Gorham, D.R.: The brief psychiatric rating scale. *Psychol. Rep.* **10**, 799–812 (1962)
25. Rajasekaran, M.P., Radhakrishnan, S., Subbaraj, P.: Elderly patient monitoring system using a wireless sensor network. *Telemed. J. E. Health* **15**(1), 73–79 (2009)
26. Roccaforte, W., Burke, W., et al.: Validation of a telephone version of the mini-mental state examination. *J. Am. Geriatr. Soc.* **40**, 697–702 (1992)
27. Rosenblum, S., Werner, P., Dekel, T., et al.: Handwriting process variables among elderly people with mild major depressive disorder: a preliminary study. *Aging Clin. Exp. Res.* **22**(2), 141–147 (2010)
28. Ruthorakuhan, M., Luedke, A., et al.: Use of physical and intellectual activities and socialization in the management of cognitive decline of aging and dementia: a review. *J. Aging Res.* (2012)
29. Saliba, D., Elliott, M., Rubenstein, L.Z., et al.: The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. *J. Am. Geriatr. Soc.* **49**(12), 1691–1699 (2001)
30. Stuck, A., Iliffe, S.: Comprehensive geriatric assessment for older adults. *BMJ* **343**, d6799 (2011)
31. The polish chamber of physicians and dentists december (2014). www.nil.org.pl/data/assets/pdf_file/0011/98930/Zestawienie-nr-04.pdf. Accessed 10 2015
32. van den Berg, N., Schumann, M., Kraft, K., Hoffmann, W.: Telemedicine and telecare for older patients—a systematic review. *Maturitas* **73**(2), 94–114 (2012)
33. Welsh, K.A., Breiter, J.C., Magruder-Habib, K.M.: Detection of dementia in the elderly using telephone screening of cognitive status. *Cogn. Behav. Neurol.* **6**(2), 103–110 (1993)
34. Yesavage, J.A.: Geriatric depression scale. <https://web.stanford.edu/~yesavage/GDS.html>. Accessed 02 2015
35. Yesavage, J.A., Brink, T.L., Rose, T.L., et al.: Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiatr. Res.* **17**(1), 37–49 (1982)

Emotion Recognition Based on Physiological Sensor Data Using Codebook Approach

Kimiaki Shirahama and Marcin Grzegorzek

Abstract This paper addresses emotion recognition based on physiological signal sequences (e.g., blood pressure, galvanic skin response and respiration) that can be obtained using state-of-the-art wearable sensors. We formulate this as a machine learning problem to distinguish sequences labelled with a certain emotion from the other sequences. In particular, we explore how to extract a feature that effectively characterises a sequence and yields accurate emotion recognition. With respect to this, existing methods rely on *hand-crafted* features that are manually defined based on prior knowledge about physiological signals. However, in addition to intensive labour, it is difficult to manually design features which can represent the details of a sequence. To overcome this, we propose a *codebook approach* where a sequence is represented with a feature describing the distribution of characteristic subsequences, called *codewords*. These are statistically justified because they are obtained by clustering a large number of subsequences. In addition, the details of the sequence can be maintained by considering the distribution of hundreds of codewords. Experimental results validate the effectiveness of our codebook-based emotion recognition method.

Keywords Ambient assisted living · Emotion recognition · Physiological sensor data · Codebook approach · Sequence classification

1 Introduction

Due to the increase of aging population, one of the most important and urgent social demands is the development of an Ambient Assisted Living (AAL) system which can support the independent life and healthcare of an elderly person [14, 17]. In this context, we are working within the research project *Cognitive Village: Adaptively*

K. Shirahama (✉) · M. Grzegorzek
Pattern Recognition Group, University of Siegen,
Hoelderlinstr. 3, 57076 Siegen, Germany
e-mail: kimiaki.shirahama@uni-siegen.de

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_3

Learning Technical Support System for Elderly funded by the German Federal Ministry of Education and Research. By focusing on the recent advancement of computers and sensor networks, we are especially developing an AAL system that monitors the status of an elderly person by recognising his/her activities based on data continuously obtained from wearable sensors together with sensors embedded in the environment.

This paper presents our progress on mental activity (i.e., emotion) recognition based on physiological signals like blood volume pressure, galvanic skin response, and respiration rate. Apart from facial, vocal and gestural expressions that can be easily controlled, it is difficult for a person to intentionally control physiological responses [10, 15]. Hence, physiological signals are reliable cues for emotion recognition. In addition, it has recently become possible to easily obtain physiological signals using wearable sensors (wristbands) [5, 12, 13]. This enables us to continuously recognise the person's emotions for a long time. Then, temporal transitions of recognised emotions are examined to evaluate his/her mental status. For example, very frequent and rapid transitions of emotions should indicate a mental disease, and a high stress is implied if the emotion *anger* is being recognised for a long time.

In Fig. 1, emotion recognition based on physiological signals is formulated as sequence classification. This aims to build a classifier that takes as input a sequence consisting of physiological signal values (e.g., respiration rates in Fig. 1), and predicts the corresponding emotion class (e.g., *anger*, *joy* or *no-emotion* in Fig. 1). Such a classifier is generally constructed in the machine learning framework. As shown in the left side of Fig. 1, *training sequences* that are already annotated with emotion classes are firstly analysed so that the classifier can capture characteristics of sequences in each emotion class. Then, it is used to categorise *test sequences* for which emotion classes are unknown.

The formulation described above involves many problems such as the *sensitivity* that physiological signal values significantly vary depending on persons and situations, and the *difficulty* of objective and precise annotation of training sequences

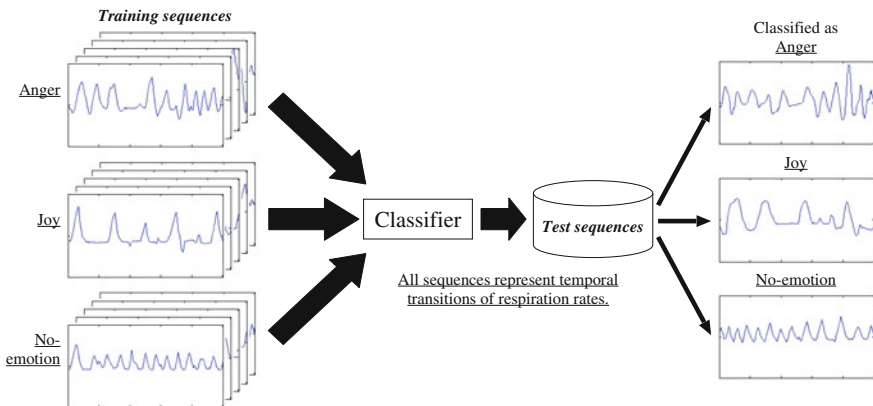


Fig. 1 An illustration of formulating emotion recognition as sequence classification

with emotion classes [8, 10, 16]. While acknowledging the importance of these problems, this paper addresses another problem of how to represent a sequence. A representation of the sequence is called a *feature*. It should express characteristics underlying the sequence, so that it becomes easier to appropriately predict the emotion class. Regarding this, most of existing methods use *hand-crafted* features such as the mean of values in a sequence, the standard deviation of values, the mean of first-order derivatives, and power spectrum in a certain frequency domain [8, 10, 11, 15, 16, 21]. However, these hand-crafted features are not statistically validated, for instance, the *mean plus 1.5* may lead to more accurate result than the mean. In addition, it is difficult to hand-craft a feature that can represent the detailed information in a sequence. Although using many hand-crafted features is one option to describe details, designing each feature requires both of prior knowledge and intensive labour [1].

Instead of hand-crafted features, we investigate *feature learning* which analyses a large amount of sequences to extract useful features [1]. As a testbed study, we adopt a simple approach, called *codebook approach*, which has been originally developed in the fields of object detection/recognition and image/video classification [9, 22]. The main idea is to firstly collect a large number of small image regions and group them into clusters of visually similar regions. Each cluster centre is called *codeword* (or visual word) and indicates a statistically distinctive small image region (a codebook is a set of such codewords). Then, an image is represented with a feature showing the distribution of codewords. Here, one important point is to use hundreds/thousands of codewords and represent the image as their combination, so that the feature can preserve the detailed information of the image. We apply the codebook approach to sequence classification by regarding small image regions as subsequences collected by sliding a window over sequences. In other words, a codebook consists of abundant codewords representing characteristic subsequences, and the details of a sequence are encoded by a feature signifying their distribution. Based on such features, a classifier is constructed to discriminate among sequences of different emotion classes.

2 Codebook-Based Emotion Recognition

Figure 2 shows an overview of our codebook-based emotion recognition method, consisting of the following three steps: Fig. 2a illustrates the first codebook construction step to group subsequences collected from sequences into clusters of similar subsequences. The centre of each cluster is regarded as a codeword. Figure 2b sketches the second codeword assignment step where the feature of a sequence is extracted by assigning each subsequence to the most similar codeword. In other words, this feature is a histogram representing the frequency of each codeword. The last classifier training/test step is based on the fact that sequences represented by such features can be considered as points in the multi-dimensional space, as seen from Fig. 2c. A classifier depicted by the dashed line is trained so as to distinguish training sequences annotated with a certain emotion class from the others. Finally,

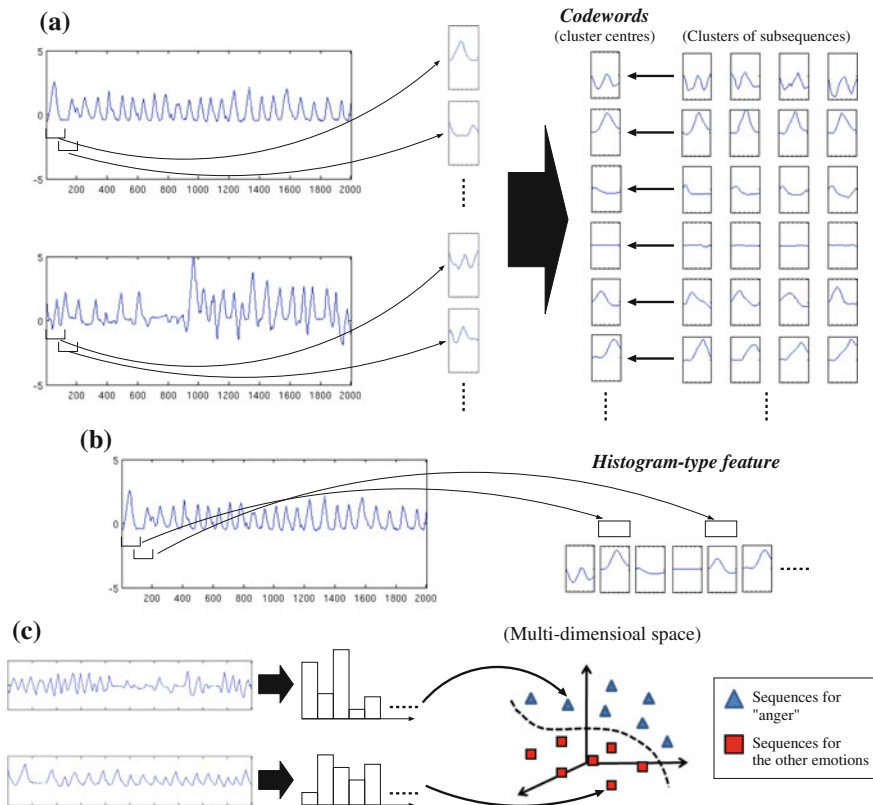


Fig. 2 An overview of our codebook-based emotion recognition method. **a** Codebook construction, **b** codeword assignment, **c** classifier training/test

the classifier determines the emotion class of a test sequence based on its location in the multi-dimensional space.

Codebook construction Let us assume a window with the size w . From each sequence, we collect subsequences by locating this window at every l time points. Then, k -means clustering [6] is conducted to find N clusters consisting of similar subsequences. The clustering repeats updating the centre of each cluster by taking the mean of subsequences belonging to this cluster, and assigning each subsequence to the cluster with the most similar centre. This repetition continues until cluster centres converge. The following two points deserve attention: First, each subsequence represents values at w time points, so it can be considered as a w -dimensional vector. Based on this, the similarity between two subsequences is measured as their Euclidean distance. Second, to achieve clustering focusing on shapes of subsequences, each subsequence is translated (normalised) so that the value at the first time point is zero. As a result of the clustering, we obtain a codebook consisting of N codewords each of which is the centre of a cluster. The right side of Fig. 2a presents an excerpt

of the codebook with $N = 64$ codewords, which are obtained by clustering 4800 subsequences collected in the setting $w = 128$ and $l = 64$.

Codeword assignment Given a sequence, we extract a histogram-type feature representing the distribution of N codewords. First, subsequences are collected from the sequence in the same way to codebook construction. Then, for each subsequence, the most similar codeword is found and its frequency is incremented. Finally, the probabilistic feature representation is obtained by normalising the frequency of each codeword so that the sum of frequencies is one.

The above approach deterministically assigns a subsequence only to a single codeword. This *hard assignment* lacks the flexibility to deal with the uncertainty in codeword assignment. For example, in Fig. 2b, the first subsequence is the most similar to the second codeword from the left, but it is also similar to the rightmost codeword. Hence, it is not reasonable to only increment the frequency of the former codeword. To handle the uncertainty, we adopt the *soft assignment* approach that implements smooth assignment of a subsequence to multiple codewords based on kernel density estimation [23]. Let x_s and c_n be the s th subsequence ($1 \leq s \leq S$) and the n th codeword ($1 \leq n \leq N$), respectively. The smoothed frequency $F(c_n)$ is computed as follows:

$$F(c_n) = \frac{1}{S} \sum_{s=1}^S \frac{K_\sigma(D(x_s, c_n))}{\sum_{n'=1}^N K_\sigma(D(x_s, c_{n'}))}, \quad (1)$$

where

$$K_\sigma(D(x_s, c_n)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{D(x_s, c_n)^2}{2\sigma^2}\right). \quad (2)$$

Here, $D(x_s, c_n)$ is the Euclidian distance between x_s and c_n , and $K_\sigma(D(x_s, c_n))$ is its Gaussian kernel value where σ is the parameter to control the smoothness. A large σ causes a strong smoothness so that smoothed frequencies of all codewords become similar. Note that when x_s is similar to c_n (i.e., $D(x_s, c_n)$ is small), $K_\sigma(D(x_s, c_n))$ is large, so x_s offers a large contribution to $F(c_n)$ (this contribution is normalised by the sum of kernel values to all the codewords, as shown in the denominator of Eq. (1)). This way, soft assignment yields a feature representing the smoothed distribution of codewords based on their similarities to subsequences.

Classifier training/test Let us consider training a binary classifier that distinguishes training sequences labelled with a target emotion class from the others. We call the former and latter *positive sequences* and *negative sequences*, respectively. To gain the high discrimination power of the classifier, we need to consider a variety of characteristic subsequences using hundreds of codewords (i.e., N is large). This means that each sequence is represented with a high-dimensional feature. Hence, a Support Vector Machine (SVM) is used because of its effectiveness for high-dimensional

data [24]. Assuming each sequence is a point in the N -dimensional space, the SVM constructs a classification boundary based on the *margin maximization* principle so that the boundary is placed in the middle between positive and negative sequences. This makes the generalisation error of the SVM independent of the number of dimensions, if this number is sufficiently large [24]. In addition, the combination of a codebook approach and SVM has been justified in the field of multimedia processing where images/videos are represented with high-dimensional features [9, 22].

Given a test sequence, the SVM checks its location with respect to the classification boundary, and outputs the degree of how likely it belongs to the target emotion class [2]. Finally, in the case where a test sequence needs to be classified into one of E motion classes, we build E SVMs each built as a binary classifier based on the above-mentioned approach. Then, we assign the test sequence to the emotion class for which the output of the corresponding SVM is the highest.

3 Experimental Results

We use the dataset introduced in [15]. It includes the following four types of physiological sequences, Blood Volume Pressure (BVP), Galvanic Skin Response (GSR), RESpiration (RES) and ElectroMyoGram (EMG). A BVP value represents an amount of blood measured by the light reflected from the skin (photoplethysmograph), a GSR value indicates a level of sweat gland activity that causes changes in the skin conductance, a RES value describes a chest cavity expansion generated by the inhalation and exhalation, and an EMG value expresses the activity of a muscle based on its surface voltages (see [7, 15] for more detail). These four types of sequences are collected for eight emotions of one person for 20 days. Those emotions are *no-emotion*, *anger*, *hate*, *grief*, *p(latonic)-love*, *r(omantic)-love*, *joy* and *reverence*. For one day, BVP, GSR, RES and EMG sequences are synchronised with each other and have values at 2001 time points sampled at 20 Hz. Figure 3 presents the BVP, GSR, RES and EMG sequences when the person felt *anger* at the first day. That is, for each type (i.e., BVP, GSR, RES or EMG), we have 160 sequences (20 days \times 8 emotions). Here, to reduce the bias of values depending on days, sequences for one day are normalised so as to have the mean zero and the variance one. However, even these normalised sequences have complex waveforms as shown in Fig. 3, so it is difficult to design hand-crafted features that can capture detailed characteristics. This implies the necessity of feature learning like our codebook approach to extract useful features for accurate classification of complex waveforms.

Using subsequences collected from 160 sequences for each type, a codebook is constructed. Since k -means clustering depends on initial cluster centres that are randomly determined, we conduct it 10 times and select the best result that yields the minimum sum of Euclidian distances between subsequences and their assigned cluster centres. Afterwards, we evaluate the performance of emotion recognition in the same way as [15]. Specifically, leave-one-out cross validation is conducted where 159 sequences are used for training and the remaining one is used for test. This is

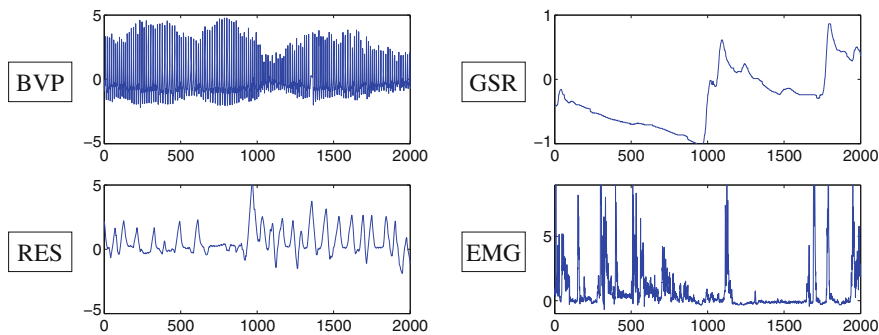


Fig. 3 Four types of physiological sequences for *anger* at the first day. Note that these sequences are normalised so as to have the mean zero and the variance one. In addition, to emphasise value changes, each type of sequence is plotted using the *vertical axis* with a different range

iterated until all the sequences are tested. Finally, the performance is measured as the accuracy (percentage) of how many sequences are correctly classified.

Parameter tuning First of all, we tune three main parameters in our codebook-based method, that is, the window size w and the sliding size l for subsequence collection, and the number of codewords N . Figure 4 shows results obtained by different configurations of w , l and N for each type of physiological signals. Here, the feature of each sequence is extracted with the hard assignment approach. The bar graph in Fig. 4a presents the distribution of accuracies on RES by setting $w \in \{8, 16, 32, 64, 128\}$ and $N \in \{64, 128, 256, 512\}$ where s is always defined as the half of w . Since it is difficult to grasp an overall trend from this distribution, we summarise it into a box plot as shown in Fig. 4b. It presents the maximum, upper quartile, median, lower quartile and minimum of accuracies. For each type of physiological signal, the left box plot summarises the distribution of accuracies using l defined as the half of w , while the right one outlines accuracies by consistently setting l to 8. As can be seen from median and maximum accuracies in Fig. 4b, $l = 8$ generally yields more accurate results than $l = w/2$. Moreover, we find that the best accuracy for each type is accomplished by $w = 64$ or $w = 128$. Thus, it can be said that, compared to $l = w/2$, more accurate results are achieved by densely collecting subsequences with $l = 8$. Regarding N , we observe that $N = 512$ stably lead to high accuracies, as exemplified in Fig. 4a. Note that, compared to w and N in Fig. 4, our preliminary experiment showed that larger w ($\in \{256, 512, 1024\}$) and N (1024) cause the performance degradation.

Overall, Fig. 4b indicates that RES is the most discriminative signal for emotion recognition, and its performance is significantly higher than the others. But, in what follows, we will show that the performance of RES is significantly improved by combining it with the other physiological signals. Finally, to avoid under-/over-estimating emotion recognition accuracies, we will continue to use the box plot representation.

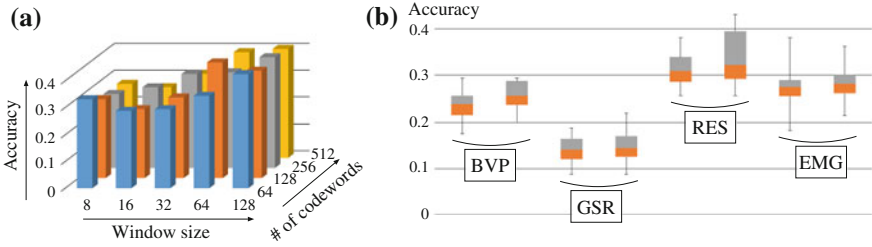


Fig. 4 Accuracies of our method depending on different parameter settings. **a** Example distribution of accuracies on RES ($l = w/2$), **b** box plots of distributions of accuracies

Before moving to the next experiment, we examine what kind of codewords are extracted by our method. For each type, Fig. 5 visualises 20 codewords extracted by the parameters tuned above (i.e., $w = 128$, $l = 8$ and $N = 512$). Small rectangular regions display codewords (subsequences) representing values at 128 time points. These codewords are centres of $N = 512$ clusters which are found on 39840 subsequences collected by $w = 128$ and $l = 8$. The number under a small rectangular region indicates the ID of the corresponding codeword. For example, the top-left small rectangular region in BVP presents the 461th codeword. As can be seen from Fig. 5, our method could successfully extract codewords as subsequences exhibiting various characteristic changes of values. A feature based on these codewords can preserve the detailed information of a sequence, while it is difficult or impossible to manually design such codewords. In addition, one can further investigate codewords to discover knowledge about relations between emotions and physiological signals. For example, codewords specific to certain emotions can be found by checking the entropy of each codeword over emotions. In Fig. 5, the 461th codeword for BVP appears 26 times over the whole set of sequences, but 25 appearances occur in sequences for *no-emotion* (entropy = 0.235). Moreover, 27 of 31 appearances of the 74th codeword for RES occur in sequences for *grief* (entropy = 0.659). This way, our method offers an easy-to-use application for physiological signal analysis, where the parameters that a user has to tune are only w , l and N .

Effectiveness of soft assignment As shown in Fig. 5, several codewords are very similar to each other, so it is rigid to use the hard assignment approach that assigns a subsequence only to the most similar codeword. Now, we examine the effectiveness of the soft assignment approach that assigns a subsequence to multiple codewords based on their similarities. Figure 6 presents box plots showing accuracies obtained by different σ s. For each type, the leftmost box plot indicates accuracies based on hard assignment (i.e., the right box plot in Fig. 4). By checking median accuracies in Fig. 6, it can be seen that soft assignment yields more accurate emotion recognition than hard assignment, although accuracies depend on σ . Based on Fig. 6, the following experiments are conducted using $\sigma = 1, 0.25, 1$ and 0.25 for BVP, GSR, RES and EMG, respectively.

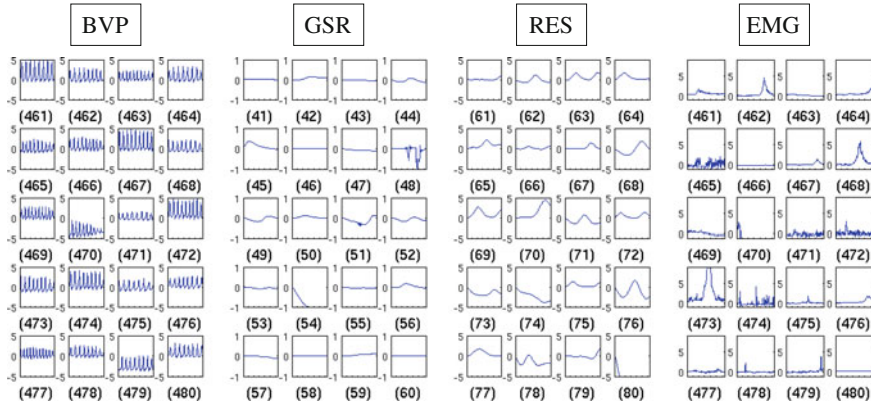


Fig. 5 An excerpt of codewords extracted with the parameters $w = 128, l = 8$ and $N = 512$

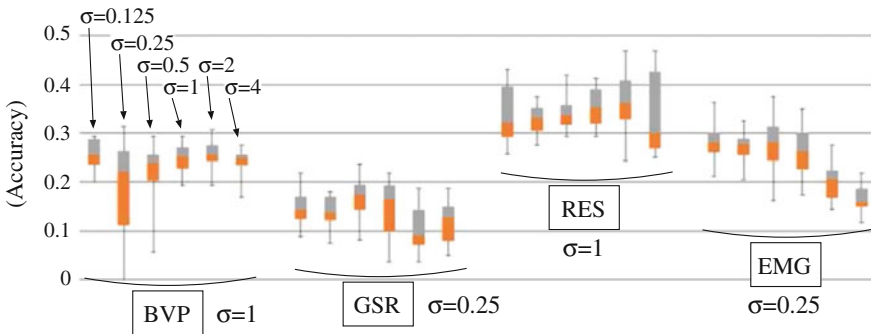


Fig. 6 Accuracies of the soft assignment approach with different σ s

Fusion of different sensor data For accurate emotion recognition, we integrate four types of sequences using the *early fusion* approach [20]. For each of 160 emotion instances (20 days \times 8 emotions), we have four features on BVP, GSR, RES and EMG sequences. The early fusion approach simply concatenates these features to create a single feature. That is, each feature has N dimensions, so the concatenation generates a $4N$ -dimensional feature. Then, SVMs are trained and tested by representing each emotion instance with such a concatenated feature. Figure 7 shows the comparison between accuracies based on single features (i.e., box plots in Fig. 6 for the selected σ s) and accuracies based on the early fusion approach. Note that it is possible to choose the best parameters for each feature and fuse such features. But, choosing the best parameters is labour-intensive and time-consuming. Hence, for the sake of simplicity and generality, we apply the early fusion approach to features all of which are extracted using the same parameters w, l and N (only σ for soft assignment is different for each feature). This enables us to draw the box plot of the early fusion approach using the aforementioned combinations of w and N ($l = 8$).

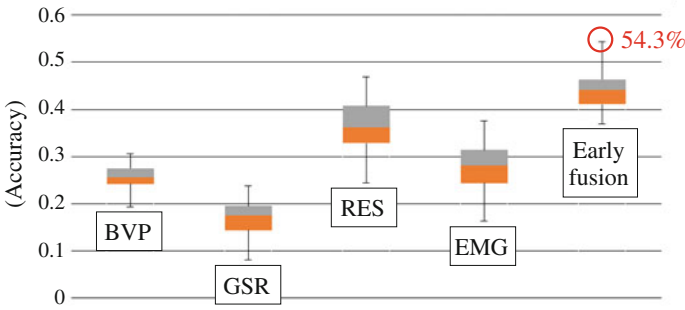


Fig. 7 Performance comparison between single features and early fusion

Table 1 Comparison between our codebook-based method and other methods

	Codebook-based	Picard et al. [17]	Raw early fusion
Accuracy	54.3 % (55.6 %)	37.5 % (40.0–46.3 %)	38.5 %

As seen from Fig. 7, the early fusion approach significantly improves the emotion recognition performance. In particular, the setting of $w = 128$ and $N = 512$ leads to the maximum recognition accuracy 54.3%.

Comparison to other methods Finally, using Table 1, we compare the performance of our codebook-based method to those of other methods. First, the performance of our method is represented with the maximum accuracy obtained by the early fusion approach. The accuracy in the parentheses is acquired by fusing five features, including four features described before, and one feature based on the codebook extracted by performing k -means clustering on amplitude spectra of subsequences in RES sequences. Each amplitude spectrum is obtained by applying FFT (Fast Fourier Transform) to a subsequence. In other words, the feature represents the distribution of characteristic amplitude spectra in a RES sequence. Note that both of the accuracies are obtained with the same parameter setting of $w = 128$, $l = 8$ and $N = 512$. Thus, these accuracies are not obtained by accident, but achieved using the most effective parameter setting.

The third column shows the accuracy based on 24 features proposed in [15]. Specifically, for each of BVP, GSR, RES and EMG sequences, six features representing the statistical information like mean, standard deviation and mean of first-order derivatives, are extracted. Then, the early fusion approach is used to fuse six features for four different types into a 24-dimensional vector. Note that our main objective is to examine the usefulness of features extracted by the codebook approach. Thus, the accuracy outside the parentheses is obtained using the same classifier as ours (i.e., SVM). On the other hand, the accuracies inside the parentheses are reported in [15] using a different classifier. The last column presents the accuracy where BVP, GSR, RES and EMG sequences containing values at 2001 time points are considered as 2001-dimensional vectors. Then, the early fusion approach is used to concatenate

these vectors into a 8004-dimensional vector. Similar to the third column, the same classifier as ours is used except features of sequences. Although one may think that the above approach which considers a raw sequence as a high-dimensional vector is very simple (very low-level), it is reported to work surprisingly well [4]. As shown in Table 1, our codebook-based method significantly outperforms the comparison methods. This validates the effectiveness of the codebook approach for extracting useful features for accurate emotion recognition.

4 Conclusion and Future Work

In this paper, we introduced an emotion recognition method based on the codebook approach. It firstly collects subsequences by sliding a window over sequences, and group them into clusters. Each cluster centre representing a characteristic subsequence is regarded as a codeword. Then, by assigning subsequences in a sequence to codewords, we extract a feature representing the distribution of codewords. In particular, to deal with the uncertainty in codeword assignment, the soft assignment approach based on kernel density estimation is adopted to obtain a feature representing a smoothed distribution of codewords. Using such features, an SVM is constructed to distinguish sequences labelled with a certain emotion class from the others. Experimental results show not only the effectiveness of features extracted by the codebook approach but also the usefulness of the soft assignment and feature fusion approaches.

Regarding our future work, a histogram-type feature used in this paper is restrictive in the sense that it only represents the distribution of pre-defined codewords, and the variation of codewords is not represented. Thus, we will investigate more sophisticated feature representations [18]. For example, we will consider the GMM (Gaussian Mixture Model) super-vector representation where codewords suitable for each sequence is estimated using Maximum A Posteriori (MAP) adaptation [19]. In addition, a variation of codewords is represented with the variance of each mixture component. Also, we will test the Fisher kernel approach which does not directly represent the distribution of codewords, but represents its first and second order differences to the reference distribution [3]. This is useful not only for considering the variation of codewords, but also for creating a very high-dimensional representation, so that a simple and fast classifier (typically linear SVM) is enough for accurate recognition.

Acknowledgments Research and development activities leading to this article have been supported by the German Federal Ministry of Education and Research within the project *Cognitive Village: Adaptively Learning Technical Support System for Elderly* (Grant Number: 16SV7223K).

References

1. Bengio, Y., Courville, A., Vincent, P.: Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(8), 1798–1828 (2013)
2. Chang, C.C., Lin, C.J.: Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**(3), 27:1–27:27 (2011)
3. Chatfield, K., Lempitsky, V., Vedaldi, A., Zisserman, A.: The devil is in the details: an evaluation of recent feature encoding methods. In: *Proceedings of the 22th British Machine Vision Conference (BMVC 2011)*, pp. 76.1–76.12 (2011)
4. Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., Batista, G.: The UCR time series classification archive (2015). www.cs.ucr.edu/~eamonn/time_series_data/. Accessed 24 Feb 2016
5. Garbarino, M. et al.: Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In: *Proceedings of the 2014 EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth 2014)*, pp. 39–42 (2014)
6. Han, J., Kamber, M., Pei, J.: *Data Mining: Concepts and Techniques*, 3rd edn. Morgan Kaufmann (2011)
7. Healey, J.A.: *Wearable and Automotive Systems for Affect Recognition from Physiology*. Ph.D. thesis, Massachusetts Institute of Technology (2000)
8. Hong, J.H., Ramos, J., Dey, A.K.: Understanding physiological responses to stressors during physical activity. In: *Proceedings of the 14th ACM International Conference on Ubiquitous Computing (UbiComp 2012)*, pp. 270–279 (2012)
9. Jiang, Y.G., Yang, J., Ngo, C.W., Hauptmann, A.G.: Representations of keypoint-based semantic concept detection: a comprehensive study. *IEEE Trans. Multimed.* **12**(1), 42–53 (2010)
10. Kim, J., Andre, E.: Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(12), 2067–2083 (2008)
11. Koelstra, S., Mühl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: DEAP: a database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* **3**(1), 18–31 (2012)
12. Metz, R.: This fitness wristband wants to play doctor. *MIT Technol. Rev.* (2014). <https://www.technologyreview.com/s/524376/this-fitness-wristband-wants-to-play-doctor/>. Accessed 24 Feb 2016
13. Microsoft Corporation: Microsoft Band Official Site. <https://www.microsoft.com/microsoft-band/en-us>. Accessed 24 Feb 2016
14. Pentland, A., Lazer, D., Brewer, D., Heibeck, T.: Using reality mining to improve public health and medicine. In: Bushko, R.G. (ed.) *Strategy for the Future of Health*, pp. 93–102. IOS press (2009)
15. Picard, R.W., Vyzas, E., Healey, J.: Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(10), 1175–1191 (2001)
16. Plarre, K. et al.: Continuous inference of psychological stress from sensory measurements collected in the natural environment. In: *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN 2011)*, pp. 97–108 (2011)
17. Rashidi, P., Mihailidis, A.: A survey on ambient-assisted living tools for older adults. *IEEE J. Biomed. Health Inf.* **17**(3), 579–590 (2013)
18. Shirahama, K., Grzegorzec, M.: Towards large-scale multimedia retrieval enriched by knowledge about human interpretation: retrospective survey. *Multimed. Tools Appl.* **75**(1), 297–331 (2016)
19. Shirahama, K., Uehara, K.: Kobe university and Muroran institute of technology at TRECVID 2012 semantic indexing task. In: *Proceedings of the TREC Video Retrieval Evaluation Workshop (TRECVID 2012)*, pp. 239–247 (2012)
20. Snoek, C.G.M., Worring, M., Smeulders, A.W.M.: Early versus late fusion in semantic video analysis. In: *Proceedings of the 13th Annual ACM International Conference on Multimedia (MM 2005)*, pp. 399–402 (2005)

21. Soleymani, M., Lichtenauer, J., Pun, T., Pantic, M.: A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* **3**(1), 42–55 (2012)
22. van de Sande, K.E., Gevers, T., Snoek, C.G.: Evaluating color descriptors for object and scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(9), 1582–1596 (2010)
23. Van Gemert, J.C., Veenman, C.J., Smeulders, A.W.M., Geusebroek, J.M.: Visual word ambiguity. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(7), 1271–1283 (2010)
24. Vapnik, V.N.: *Statistical Learning Theory*. Wiley-Interscience (1998)

Human Activity Recognition Using Smartphone Sensors

Marcin D. Bugdol, Andrzej W. Mitas, Marcin Grzegorzek,
Robert Meyer and Christoph Wilhelm

Abstract In the paper a human activity recognition system has been presented based on the data gathered with the smartphone sensors. The acceleration, magnetic field and sound have been registered and four different activities of daily living has been recognized i.e. riding a bike, driving in a car, walking and sitting. Two version of Support Vector Machine (SVM) classifier have been employed and the obtained results are promising.

Keywords Human activity recognition · Smartphones · Support vector machine

The Human Activity Recognition (HAR) systems aim to automatically determine what people do on the basis of signals recorded from different sensors. These systems can be divided into two main types: external and wearable [11].

In the first approach the sensors are placed in a fixed position so the people have to interact with the system and they are confined to a certain area. The most popular are video cameras [20]. Typically they can register a 2D images but it can also record a 3D sequences when two or more devices are employed [1, 17]. In order to track the human activity also the depths sensors (especially Time-of-Flight cameras) are introduced. The solutions are based on tracking both the whole body [7, 14] as well as only it certain parts, in particular joints [19]. Although this approach allows obtaining a full information about the human activity and his/her environment. The image processing methods are very time consuming and not resource-efficient. Moreover, many people are not willing to be monitored permanently, except for the

M.D. Bugdol (✉) · A.W. Mitas
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: marcin.bugdol@polsl.pl

M. Grzegorzek · C. Wilhelm
Pattern Recognition Group, University of Siegen,
Hoelderlinstr. 3, D-57076 Siegen, Germany

R. Meyer
Neural Information Processing Group, Technische Universität Berlin,
Marchstr. 23, D-10587 Berlin, Germany

situation where one's life or health might depend on it i.e. telecare of elderly or disabled people [12].

The wearable sensors do not have the mentioned limitation and are more popular in HAR application [11]. It is possible to register many more parameters using them. For example in [15] the temperature has been measured and in [8] the HRV activity has been monitored. A complex system for monitoring motion, location and others signals has been presented in [3, 12, 13]. However, such systems usually require specifically designed devices that have to be attached to the human body [5, 20] and may cause discomfort in daily life. Moreover, the registered data not being processed by the sensors is transferred to external devices like a laptop or PDA [11].

The smartphones become increasingly popular in the world, and their standard feature enables recording information that can be used for activity recognition and it is an emerging field of scientific studies [2, 4, 6, 9, 10, 18].

The aim of this paper is to distinguish between different types of user movement using cell phone sensors namely accelerometer, coordinate systems and the microphone. The rest of the paper is organized as follows: Sect. 1 presents the recorded data and features extraction. In Sect. 2 the experimental results of activity recognition are described and Sect. 3 will give a summary of the paper.

1 Materials and Methods

It has been decided to incorporate four different classes that will be discriminated i.e.: *ByBike*—riding a bike; *InCar*—driving in a car; *OnFoot*—walking; *TvWatching*—being (figuratively) tied to an armchair.

1.1 Acquisition Setup Set

The data has been gathered using two different mobile phones Samsung Galaxy S Plus and HTC Hero both with Android operating system. The Samsung Galaxy S Plus runs Android 2.3.6, whereas the HTC Hero runs Android 2.1. Accordingly, the API to agree on was version 7. In addition, none of the phones provide a gyroscope. Thus, the sensors like the Linear Accelerometer which already adjusts an acceleration measurement by subtracting the force of gravity or precise orientation measurements could not be considered.

The information about acceleration, the magnetic field and sound frequency has been recorded. Acceleration has been measured with the standard accelerometer and returns a three dimensional vector of forces applied along three orthogonal directions. The force is measured in m/s^2 . The magnetic field has been recorded according to the smartphone's coordinate system and measured in μT . Note that the magnetic field measures can be used in combination with the accelerometer to (very) coarsely

determine the phone’s orientation. The sound has been recorded by the phone’s microphone and the frequency was estimated by the zero-crossing method.

The smartphones have been carried in the pocket by two person. 1004 samples of 30s each have been recorded that contain 133 instances of the class *ByBike*, 425 instances of the class *InCar*, 209 instances of *OnFoot*, and 237 instances of *TvWatching*.

1.2 Features Extraction

The raw data is not suitable for classification purpose because it cannot be guaranteed that every instance encompasses sensor recordings performed in equal time intervals. In fact, most often not even the same amount of measurements have been made for each data instance. To overcome this problem a simple scheme of feature extraction is applied. The measurement period of 30s was discretized into intervals of 250 ms. This leads to a regular grid partitioning of the multidimensional time series data space. Next, it is checked along each dimension whose measurement within the raw data is the closest to a boundary of such an interval and selected as a feature vector entry (Fig. 1).

As a consequence, dividing 30s into intervals of 250 ms results in 120 feature values for each dimension of each sensor. Thus, a full feature vector exhibits 840 dimensions. The interval size of 250 ms is a compromise between the dimensionality (still less than 1000) and a frequent consideration of sensor measurements, that is 4 per second. The features were normalized and then reduced using Principal Component Analysis (PCA) method.

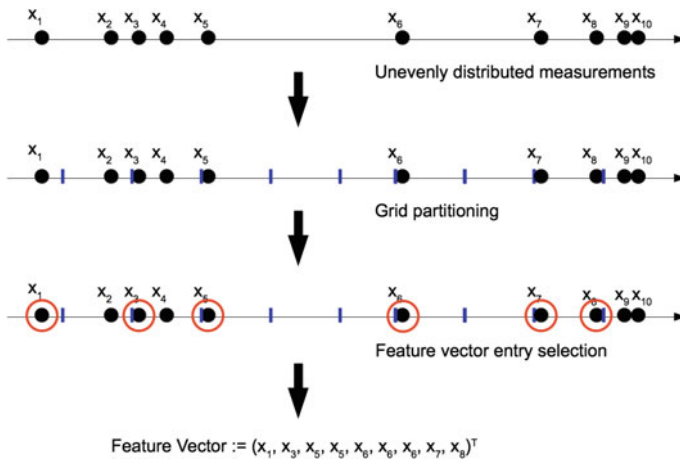


Fig. 1 Grid partitioning of feature space of one dimension with unevenly distributed measurements

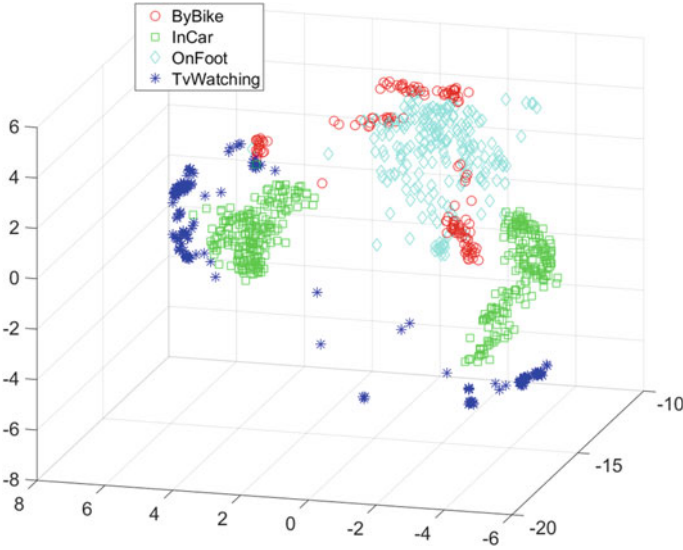


Fig. 2 The first 3 features obtained employing PCA

Figure 2 shows a 3D plot of rescaled and reduced data. It shows that linear separation of classes with such a low dimensionality is impossible. Hence, by considering a higher feature space dimension a SVMs classifier might be able to perform a successful human activity recognition.

The classification has been performed using two types of Support Vector Machine [16]. The linear approach has been followed by the non-linear SVM with radial base function (RBF) kernel. To quantify the performance of this classifier a 10-fold cross validation was used.

In the first attempt the linear SVM has been applied. The optimal number of PCA dimensions n and the value of the SVM cost parameter C have been calculated. A full factorial design and all pairwise combinations of values given below have been tested. The parameter n varies as $n \in \{2^2, 2^3, 2^4, 2^5, \dots, 2^9\}$ and the full non PCA processed number of dimensions equals 840. Moreover, C is varied as $C \in \{2^{11}, 2^9, 2^7, \dots, 2^5\}$. Using exponentially growing parameter values ensures that a relatively wide area of the parameter space has been sampled.

To achieve better results the parameter space has been explored further with a focus on values around $C = 8$ and $n = 8$. For the second run the n varies as $n \in \{5, 6, 7, \dots, 15\}$ and C as $C \in \{2^2, 2^{2.25}, 2^{2.5}, \dots, 2^6\}$.

In order to improve the classification performance the RBF kernel for non-linear classification has been applied. The parameter C has been set to $2^{3.25}$ and n to 8. The parameter γ , which determines the width of the radial basis function, is varied as $\gamma \in \{2^{-11}, 2^{-10.75}, 2^{-10.5}, \dots, 2^{2.5}\}$.

2 Results

The correlation between the obtained principal components and the input features has been analysed. The results indicates a very low contribution of parameters related to the acceleration along axis z and the magnetic field along the same axis to the new feature vector. The others input features are highly correlated and with the new feature vector.

In Table 1 the best three results for the linear SVM are presented. The best configuration uses a cost parameter of $C = 8$ and $n = 8$ dimensions. In this case more than 97 % of the test patterns were classified correctly.

The best results for the second trial of the SVM are presented in Table 2. The second run does not offer pertinent results and indicate that the linear SVM may not be employed in further application. Moreover, both runs showed that the choice of C has a weak impact on the performance whereas the number of dimensions is a crucial parameter. So the optimal number of dimensions appears to be 8.

The best results for the non-linear SVM are given in Table 3. More than 99.5 % of the test data has been correctly classified on average and no mistakes whatsoever have been made on the training sets. In fact, the best non-linear SVM ($C = 2^{3.25}$, $n = 8$, $\gamma = 2^{-2.25}$) significantly outperforms its linear counterpart in terms of average test error. This has been verified with a non-parametric, non-directional Wilcoxon Rank Sum test [$N = 100$, $p < 0.001$].

Nonetheless, the non-linear SVMs need many more support vectors to specify the classification manifold/hyperplane than the linear ones.

Table 1 The three best configurations in terms of average test error are shown

C	8	32	2
n	8	8	8
Support vectors	90.37 (± 4.76)	89.01 (± 4.90)	95.33 (± 4.72)
Training error (%)	2.00 (± 0.21)	1.93 (± 0.21)	2.06 (± 0.20)
Test error (%)	2.47 (± 1.60)	2.50 (± 1.60)	2.55 (± 1.67)

The standard deviations are given in brackets

Table 2 The three best configurations in terms of average test error are shown

C	$2^{3.25}$	2^6	$2^{3.75}$
n	8	8	8
Support vectors	89.90 (± 4.93)	88.32 (± 4.87)	89.45 (± 4.94)
Training error (%)	1.96 (± 0.22)	1.93 (± 0.22)	2.95 (± 0.22)
Test error (%)	2.49 (± 1.44)	2.50 (± 1.42)	2.50 (± 1.42)

The standard deviations are given in brackets

Table 3 The three best configurations in terms of average test error are shown ($C = 2^{3.25}$, $n = 8$, for all configurations)

γ	$2^{-2.25}$	2^{-2}	$2^{-2.5}$
Support vectors	191.87 (± 4.05)	209.04 (± 4.23)	178.41 (± 3.80)
Training error (%)	0 (± 0)	0 (± 0)	0 (± 0)
Test error (%)	0.29 (± 0.59)	0.31 (± 0.56)	0.39 (± 0.62)

The standard deviations are given in brackets

3 Summary and Future Works

A classification system of 4 different types of movement based on sensor recordings from smartphones was presented. A rather simple discretization method to extract very large feature vectors from the raw data as well as an approach to reduce the feature vector dimensionality via principle components analysis was described. The SVM yields a very good classification rate which indicates that the system might be successfully applied to more realistic scenarios.

Future work might encompass collecting and recording a much larger and much more representative samples of different movements. Other and more similar types of movement could be tackled as well. Furthermore, within this work only the offline classification of previously collected data was considered. Thus, incorporating on the fly classification potentially via a client-server architecture or even directly on the smartphone could also be investigated in the future.

References

1. Aggarwal, J., Xia, L.: Human activity recognition from 3D data: a review. *Pattern Recogn. Lett.* **48**, 70–80 (2014)
2. Anguita, D., Ghio, A., Oneto, L., Parra, X., Reyes-Ortiz, J.: Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In: Bravo, J., Hervás, R., Rodríguez, M. (eds.) *Ambient Assisted Living and Home Care. Lecture Notes in Computer Science*, vol. 7657, pp. 216–223. Springer, Berlin (2012)
3. Badura, P., Pietka, E., Franiel, S.: Acceleration trajectory analysis in remote gait monitoring. In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 4615–4618 (2014)
4. Choujaa, D., Dulay, N.: Tracme: Temporal activity recognition using mobile phone data. In: *IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, 2008. EUC '08*, vol. 1, pp. 119–126 (2008)
5. Costa, Â., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. *Expert Syst. Appl.* **39**(15), 12192–12204 (2012)
6. Guiry, J.J., van de Ven, P., Nelson, J., Warmerdam, L., Ripper, H.: Activity recognition with smartphone support. *Med. Eng. Phys.* **36**(6), 670–675 (2014)
7. Jalal, A., Uddin, M., Kim, J., Kim, T.S.: Daily human activity recognition using depth silhouettes and R transformation for smart home. In: Abdulrazak, B., Giroux, S., Bouchard, B., Pigot, H., Mokhtari, M. (eds.) *Toward Useful Services for Elderly and People with Disabilities. Lecture Notes in Computer Science*, vol. 6719, pp. 25–32. Springer, Berlin (2011)

8. Jatoba, L.C., Grossmann, U., Kunze, C., Ottenbacher, J., Stork, W.: Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity. In: Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, pp. 5250–5253 (2008)
9. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl.* **12**(2), 74–82 (2011). March
10. Kwon, Y., Kang, K., Bae, C.: Unsupervised learning for human activity recognition using smartphone sensors. *Expert Syst. Appl.* **41**(14), 6067–6074 (2014)
11. Lara, O., Labrador, M.: A survey on human activity recognition using wearable sensors. *Commun. Surv. Tutorials IEEE* **15**(3) 1192–1209 (2013)
12. Mitas, A.W., Rudzki, M., Skotnicka, M., Lubina, P.: Activity monitoring of the elderly for telecare systems—review. In: Pietka, E., Kawa, J., Wieclawek, W. (eds.) *Information Technologies in Biomedicine, Volume 4. Advances in Intelligent Systems and Computing*, vol. 284, pp. 125–138. Springer (2014)
13. Mitas, A., Rudzki, M., Wieclawek, W., Zarychta, P., Piwowarski, S.: Wearable system for activity monitoring of the elderly. In: Pietka, E., Kawa, J., Wieclawek, W. (eds.) *Information Technologies in Biomedicine, Volume 4. Advances in Intelligent Systems and Computing*. Springer International Publishing, vol. 284, pp. 147–160 (2014)
14. Ni, B., Wang, G., Moulin, P.: Rgbd-hudaact: a color-depth video database for human daily activity recognition. 1147–1153 (2011)
15. Parkka, J., Ermes, M., Korpiä, P., Mantyjarvi, J., Peltola, J., Korhonen, I.: Activity classification using realistic data from wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* **10**(1), 119–128 (2006). Jan
16. Scholkopf, B., Smola, A.J.: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge (2001)
17. Urtasun, R., Fua, P.: 3d tracking for gait characterization and recognition. In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings, pp. 17–22 (2004)
18. Varkey, J., Pompili, D., Walls, T.: Human motion recognition using a wireless sensor-based wearable system. *Pers. Ubiquit. Comput.* **16**(7), 897–910 (2012)
19. Wang, J., Liu, Z., Wu, Y., Yuan, J.: Mining actionlet ensemble for action recognition with depth cameras. 1290–1297 (2012)
20. Yang, C.C., Hsu, Y.L.: Remote monitoring and assessment of daily activities in the home environment. *J. Clin. Gerontol. Geriatr.* **3**(3), 97–104 (2012)

Assistive Integrated Personal Trainer System

Joanna Furmaniak, Roman Kaczorowski, Maciej Kasprzycki,
Bartosz Zgrzeba, Zygmunt Kubiak and Ewa Łukasik

Abstract The paper presents an Assistive Integrated Personal Trainer System called MKUFUNZI—a low cost solution to self and remote monitoring physical activities of users. It consists of (1) a sensory wearable monitoring device equipped with a pulse meter, an accelerometer and a gyroscope to acquire personal data for further processing, (2) a mobile application for collecting, analyzing and sending data to the external Web server and (3) a Web application for a supervisor. MKUFUNZI also provides a mini-medical patient card with recorded health parameters. It can help patients to plan times of medicines intake. The Web application provides access to user data that can be observed by the professional trainer controlling several trainees. The system is presented in the context of intelligent personal informatics model enabling, in the future, to supervise training sessions with reduced human trainer contribution.

Keywords Assistive technology · Personal informatics · Personal trainer · Mobile technology · Internet of things · Embedded systems · Elderly people

1 Introduction

The development of mobile technology, common use of mobile devices and growing popularity of wearable devices caused the expansion of so called personal informatics systems that have been defined as those that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge [8]. Personal informatics systems provide people means of investigating information about themselves and their everyday activities and has been mainly explored in

J. Furmaniak (✉) · R. Kaczorowski · M. Kasprzycki ·
B. Zgrzeba · Z. Kubiak · E. Łukasik
Institute of Computing Science, Poznań University of Technology,
Piotrowo 2, Poznań, Poland
e-mail: Joanna.Furmaniak@student.put.poznan.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_5

the context of physical activity, which is a good area for revealing factors affecting human behavior [5, 7]. There are many popular easily accessible mobile applications that support tracking physical activities of users. While the aspect of data collection is easier through mobile technologies, more advanced forms of assistance are not presently available and should be developed within the next generation personal informatics systems [5]. The same concerns systems supporting tracking physical activities of both young people and people with special needs, e.g. elderly, impaired, rehabilitated or hospitalized, to point out only a few.

The paper presents an initial solution to be used in the Assistive Home Lab for elderly or impaired people to be built in the Institute of Computing Science, Poznan University of Technology. The goal is to provide a testbed for numerous intelligent assistive devices and assistive services before introducing them into the real-life scenarios in homes of potential users (patients). The list of possible sensory devices to communicate in the Assistive Home Lab is long, from intelligent entryphone, through localization, activity, security, and nonstandard communication devices to integrated personal trainer system that will be presented in this paper. All of them may be linked together according to the Internet-of-Things paradigm.

The paper is structured as follows: Sect. 2 discusses hardware and software supporting physical activities available on the market and presents related works in the domain of modeling personal informatics systems. Section 3 presents details of the Assistive Integrated Personal Trainer, Sect. 4 discusses potential areas of MKU-FUNZI application and Sect. 5 concludes the paper.

2 Related Work

One of the most important factors to mitigate the effects of age and enabling successful aging is regular physical activity. There is no age restrictions for the practice of physical training. Regular physical training is recommended for seniors at any age, even among people over 80, as well as for people with various kinds of disability and impairment. It is observed that a growing number of people over 50 starts exercising. Therefore a tracker of physical activities and related health parameters being used by young working generation addicted to technology and spending time actively may be also used by a group of people with special needs, including elderly people and seniors [6]. The system described in this paper is presented within two perspectives: popular mobile hardware and software available on the market for tracking physical activities and a model of personal informatics system with consideration for integrating intelligent computing in the design of future personal trainer system [11].

2.1 *Mobile Hardware and Software for Tracking Physical Activities*

Hardware A common trend within a growing group of young people addicted to technology, working in companies with a young staff and flexible working time [6] is to spend time actively. Therefore, they look for ready-to use hardware solutions and program applications to monitor their fitness achievements and share them with friends. The market responds to such demands—manufacturers of mobile devices offer various wearable activity trackers monitoring and recording fitness activities. The most popular are Smart Watches (devices combining the functionality of a digital watch and a smartphone) [3] or Smart Bands (devices poorer than Smart Watches, created mainly for collecting behavioral data and displaying basic information from a smartphone). There is a multitude of mobile applications instructing how to train appropriate parts of body and supporting running activity logs. Also, the number of medical monitoring applications grows.

Smart Bands, like, e.g., Samsung Galaxy Gear Fit, LG Lifeband Touch or Sony Smart Band Talk, have got a small touch-screen display (from 0.9 to 1.8 inch), Bluetooth connectivity, pulse sensors, gyroscope, accelerometer and are water resistant. These solutions are geared primarily for physical activity monitoring. They provide training modes, enabling more accurate analysis of exercises and have a function of displaying text messages or rejecting incoming phone calls. Smart Watches constitute more sophisticated equipment as compared to Smart Bands described above. They usually have ca 2 inch touch display, a built-in processors (2 or 4 core) similar to those used in smartphones, a memory and Bluetooth or WiFi modems. It is possible to fully control these sensors via a mobile phone and make use of a part of the functionality of the phone itself. Huawei SmartWatch, Motorola 360 (2nd generation) and Samsung Gear S (R750) have sensors similar to those applied in Smart Bands (pulse sensor, gyroscope and accelerometer). Unfortunately, the greater the number of functions offered by Smart Watches, the higher are their prices. Table 1 compares functionality and approximate prices of selected devices available on market and a system MKUFUNZI proposed in this paper.

Software There is a multitude of mobile applications in the field of assisting healthy life and physical activity that motivate and coach users in tracking their progress. One of the most popular applications is Endomondo—application that tracks runs, bicycle rides, and other outdoor activities with good accuracy and a simple interface [4]. Data collected can be shared on the Facebook. Another very popular application is a 7 min Workout [1], reminding and motivating users to perform 7 min long physical exercises daily.

Much more powerful solution is offered by Samsung within the S-Health application [12]. This application is designed to work with both a smartphone as a sensor and a dedicated device, Samsung Gear Fit. It also provides a medical file that allows for the collection of data on heart rate, blood pressure or blood glucose level. Table 2 compares characteristics of selected software for mobile devices available on market and a system MKUFUNZI proposed in this paper.

Table 1 Comparison of personal activities measuring devices

Characteristics	Smart band	Smart watch	MKUFUNZI
Purpose	Sport and personal activity monitoring, basic smartphone remote-control	Almost full smartphone functionality (can work as separate device under Android Wear OS or as a smartphone remote-control)	Sport and personal activity monitoring
Separate from smartphone?	No	Yes	No
Medical and activity sensors	Yes	Yes	Yes
Connectivity	Bluetooth	Bluetooth, WiFi	Bluetooth
Display	Touch, 0.9''–1.8''	Touch, app. 2.0''	No (planned for future)
Others	Displaying text messages or e-mails, rejecting calls	Almost full smartphone functionality	No (in future displaying trainer messages)
Price	ca 100 EUR	ca 300–320 EUR	prototype: ca 40 EUR serial production: ca 20 EUR

Missing Aspects There are many other, more or less successful solutions on the market. However most of them are oriented either to support physical exercises or to conduct medical file (e.g. MS HealthVault or Apple Health [2]), but not both.

On the other hand, in the range of devices available, prices are relatively high in relation to the cost of electronic components. Their manufacturers are not open for the cooperation with non-smartphone companies and applications from independent authors can be hardly used. Thus, the authors of this paper decided to design a generic system, playing the role of an assistive personal trainer, that may be connected to a variety of dedicated mobile devices. This system supports personal physical exercises of a user that are supervised by a remote trainer and that are stored in medical records of the user, who has access to both mobile and Web applications. This assistive personal trainer may be a testbed for training data analysis towards future intelligent system.

2.2 *Intelligent Personal Informatics Model Perspective*

Data collection and visualization are the most popular means of personal data processing including physical activities. As it was stated in Sect. 1, intelligent systems are

Table 2 Comparison of selected mobile software for monitoring personal activities

Functionality	Endomondo [4]	7 min [1]	Zdrowia Nies [14]	S-Health [12]	Apple health [2]	MS HealthVault [10]	MKUFUNZI
Training (distance, repetitions, calories)	Yes	Yes	Yes	Yes	No	No	Yes
Medical file	No	No	Poor	Yes	Yes	Yes	Yes
Measuring device	Smartphone			Smartphone, SmartWatch, Smart band	Medical devices compatible with API		MKUFUNZI device

considered desirable if they can provide personalized advice to the user. Several models for understanding how people use personal informatics tools have been proposed in the literature with case studies on physical activities given as examples. These models may serve as guidelines for providing feasible solutions in future systems. Li et al. [8] have proposed a five-stage action model consisting of iterative transitions between preparation, collection, integration, reflection, and action. The model also identifies the barriers towards a given action. Those barriers are important from the point of view of a planned goal of user's behavior change.

Epstein model includes an additional stage of decision making towards tracking the activities and selecting tools, as an ongoing process of collection, integration, reflection, and lapsing of tracking that may later be resumed [5].

Ohlin and Olsson [11] have recently identified twelve key factors for integrating intelligent computing in the design of future personal informatics systems related to:

- human understanding and motivation,
- outcome focused cooperation,
- process focused cooperation,
- assistance focused cooperation.

Thus, a cycle of human-computer interaction and transitions facilitating cooperative actions have to be taken into account: human driven cooperation, revised computer system behavior, computing-driven cooperation and revised human behavior. Based on these transitions, 12 considerations have been proposed for designing personal informatics systems. Their full list can be found in [11]. We point out here only the most important for future development of our system. Since the personal informatics system is a cooperation system between human and computer also in terms of human engagement and motivation, attention has to be focused not only on models of the domain (e.g. specific type of physical exercise or biometric factors), but also on:

- models of how people use the personal informatics system itself,
- exploration of forms of real-time system-provided feedback to decrease the need for sophisticated post hoc data analysis by the user,
- ways of exploring real-time feedback in past performance, projections, and/or goals,
- exploring ways of including users in processes of automatic data collection,
- considering ways of letting users give feedback on interjections from the personal informatics system.

The above user-centric modelling issues are equally important as those related to the choice of wearable sensors, appropriate signal processing methods, activity classification or richer understanding of the tracked activities.

3 MKUFUNZI—Assistive Integrated Personal Trainer System

3.1 General Description

Mkufunzi in Suahili means *a trainer*. The proposed Integrated Personal Trainer called MKUFUNZI is a system supporting people in their physical training and collecting medical parameters records. It consists of (Fig. 1):

- **sensory wearable monitoring device** equipped with pulse meter, accelerometer and gyroscope to collect data for further processing; it communicates with the mobile device via the Bluetooth,
- **mobile application** for Android devices, connecting with training device and collecting, analyzing and sending data to the external Web server. It can also provide some kind of mini-medical patient card with recorded blood pressure, weight or pulse database and can help patients to plan times to medicines intake,
- **Web application** that provides access to user data, such as training history or mini-medical patient card. Web application provides also a supervisor (trainer/doctor) mode, which lets viewing and analyzing data of numerous trainees. It could be useful for doctors to monitor patient blood pressure, for nurses to check if patients walk enough at home and how high is their pulse.

Mobile and Web applications may be used without a mobile device and serve as a medical mini-file collecting daily pressure or pulse measurements, body mass index (BMI) or drugs organizer for planning and controlling drugs intake (only in a Web application).

But the most important and unique feature of the presented system is the introduction of a new actor to the system *a trainer*, who is a privileged user of the Web application. This innovation consists in the fact that the trainer can observe current training activities of her trainee. In addition, she has the insight into the records of the trainee. All this was introduced in order to better plan and adapt training program

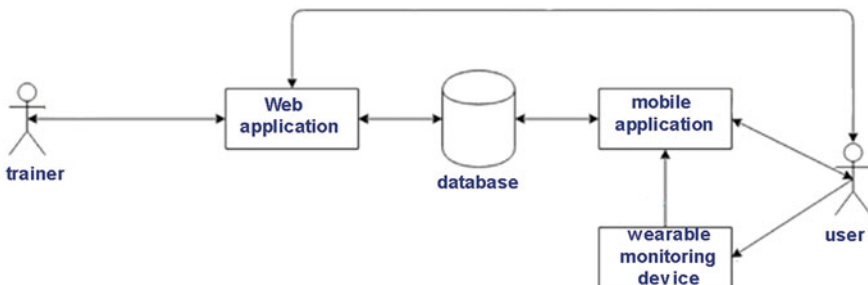


Fig. 1 Elements of the assistive integrated trainer system MKUFUNZI

to the needs of the user and to monitor health and dangerous changes in vital parameters, such as high blood pressure or high heart rate, excessive blood glucose level, etc.

3.2 Sensory Wearable Monitoring Device

Wearable computer is *any body-worn computer that is designed to provide useful services while the user is performing other tasks* [13]. The wearable monitoring device has a task of collecting health and sports parameters from the sensors and sending them to the application installed on the mobile device (a smartphone). Currently the prototype device is equipped with accelerometer/gyroscope MPU-6050, and pulse sensor Seedstudio MED03212P. Data communication is provided by Bluetooth HC-05 and all functions are controlled by a microcontroller ATmega 328p. The prototype is characterized by a fairly large size (115 × 65 × 37 mm) and a relatively simple structure (Fig. 2b). The etching of printed circuit board was done using *Do It Yourself* (DIY) thermotransfer method. Power supply is provided by four AAA batteries. This simple construction resulted a very low price of the device. The whole device is placed in a casing and is to be fastened on the wrist of the user (Fig. 2a).

In further stages of the project the miniaturization of the device is planned by the professional PCB production methods, and the change of 5 V power supply to a small battery with a capacity of min. 1200 mAh. The introduction of the device for mass production and to obtain agreements with component suppliers and companies producing PCB layouts could result in a reduction in price of the device in its current form to a level of approximately 50 PLN (ca 12 EUR). It is also planned to equip the device with a small display for receiving short messages and commands from the trainer.

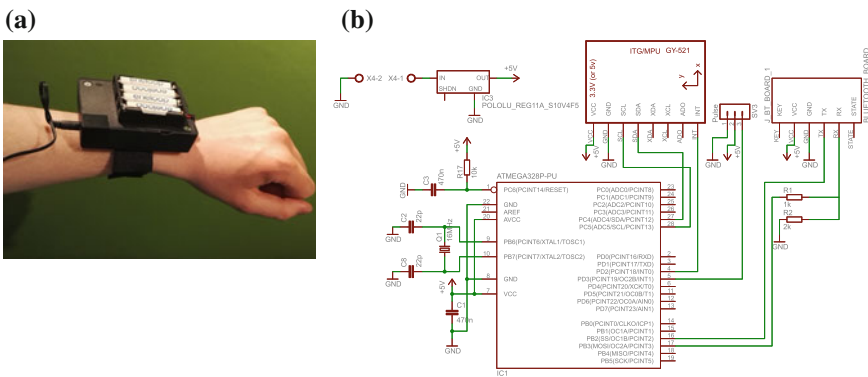


Fig. 2 Wearable monitoring device. a Photo. b Schema

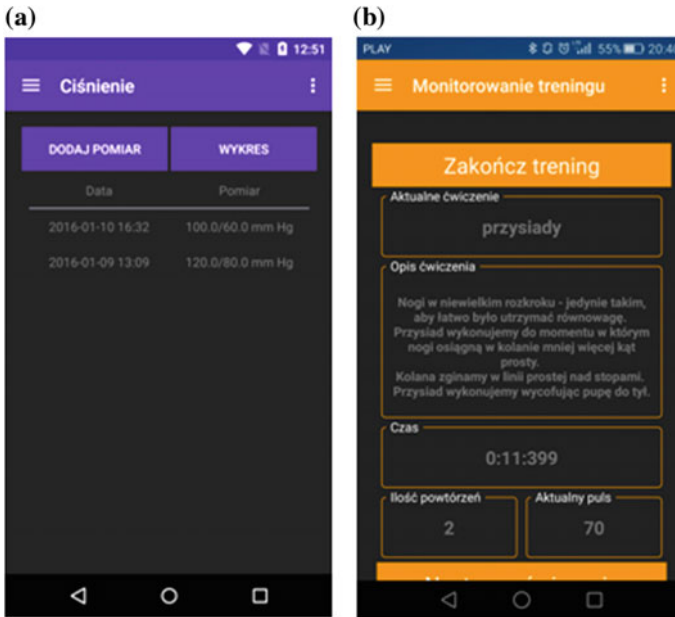


Fig. 3 Trainer view in mobile application. **a** Bloodpressure. **b** Monitoring

3.3 Mobile Application

The wearable device should communicate with the mobile application according to the paradigm *one wearable device—one application on the mobile device*. Mobile application, connecting with wearable monitoring device and collecting, analyzing, visualizing and sending data to the external Web server is thus a pivotal part of the system, linking all its components. It can also provide some kind of mini-medical patient card with recorded blood pressure, weight or pulse database and help patients to plan times to medicines intake (Fig. 3). It may also exist as a separate application, independent from training device, fulfilling a number of functions, such as e.g. communication with the server to display the data stored or exploit training history records and mini-medical files. In addition to presenting the information, it is possible to manually add data to the mini-medical record (via a keyboard entry). In order to identify the user, the application allows for logging or performing a full process of a new user registration. The user can also edit her profile data. Modern techniques and software concepts enable flexible functionality of mobile applications and their multi-device adaptation. The proposed system has been built using:

- Java with Android API allowing for adapting the application to various devices and providing multiple entry points;

- Material Design—new interface design language; Material Design is a comprehensive guide for visual, motion, and interaction design across platforms and devices [9];
- Active Android as an Object-relational mapping (ORM) tool; ORM is a programming technique for converting data between incompatible type systems in object-oriented programming languages;
- Model-view-controller (MVC) architectural pattern enabling separation of internal data representations from the user presentation layer;
- Singleton anti-pattern, that denies the use of the instantiation of a class to one object only.

3.4 Web Application

The Web application is designed for users playing three different roles: an administrator, a user (trainee) and a trainer. During registration, the user must provide information on its role in the system: a trainer or a user. An administrator is appointed by the manager of the system. A trainer can start using the Web application only after activation by the administrator while the user immediately gains access to the system. The application allows the administrator to manage trainers, to activate and deactivate their accounts, to add, delete and edit types of measurements on file, and to define sets of exercises for users (Fig. 4). A trainer assigns to herself a trainee and starts observing her results. The most interesting option for the trainer function is observation of trainees activities and their results in real time. In addition, she has the insight into the health records of a trainee and her prescribed medicines in order

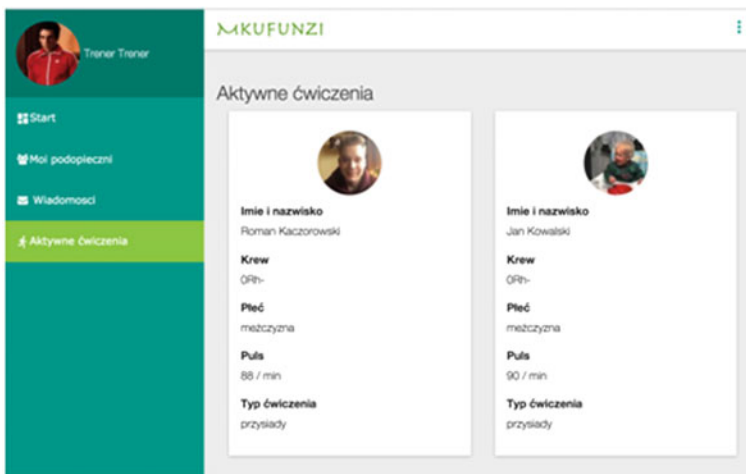


Fig. 4 Trainer view in web application

to better plan and adapt training to her needs and to monitor health and dangerous changes in vital parameters, such as high blood pressure, high heart rate, high blood glucose level etc.

A trainee, as a most active actor in the system has the access to her results and parameters as well as scheduled training exercises so that she can verify training plans fulfillment, monitor health parameters (e.g. heart rate during exercises) and medicines intake. The system allows for the text messaging in both directions using built-in communication messages.

For the purpose of tracking a training progress, as well as for collecting data for further research, the Web application enables keeping a life parameters file of trainees, as weight, height, blood pressure or body temperature. As for now data are presented in tables and in graphs, but other visualization forms may be introduced too.

Web application has been built in *Ruby on Rails* technology obeying the DRY rules (*Dont Repeat Yourself*) ensuring code transparency and using the MVC (*Model-View-Controller*) model.

3.5 Tests

The prototype of the system has been thoroughly tested. Three-phase tests have been performed: developer tests, exploratory tests of mobile/Web applications, acceptance tests for training device and complex pre-production tests of the whole system. Developer tests were performed using debugger and logging systems or just by executing all user cases. Exploratory tests were executed by a team member acting as an average user, checking the functionality and intuitiveness of the system. Correctness of the training device measurements was verified during acceptance tests. The complex pre-production system testing was accomplished during one hour long training session at the gym. Two team members used two prototypes connected to two smartphones. The third team member, acted a trainer role and checked trainer functions over the Web.

3.6 Usability

In order to collect users opinions concerning MKUFUNZI, a comparison with Apple Watch has been performed. Measurement accuracy, intuitiveness of use and a price were compared. Measurement values were almost the same for both devices and a reference. User experience was better for Apple Watch, as it is much smaller, lighter and has a touch display. However MKUFUNZI was not designed to compete with existing solutions—it was designed to fill a market gap, as it is a unique system offering a *trainer mode* letting a trainer to remotely monitor exercises of trainees

in real time. This feature is fully appreciated by users. MKUFUNZI is also much cheaper and thus may be commonly used.

3.7 Further Development of MKUFUNZI

Although the MKUFUNZI system constitutes a fully operational and useful product, there is still room for the improvement. The most important development goals for further work concern both hardware and software, as well as data modeling leading to intelligent version of the system. Planned development include:

- miniaturization of the monitoring device to make it competitive with solutions existing on the market;
- replacing Bluetooth HC-05 device by LE 4.0;
- adding security module for Bluetooth communication to assure data privacy of the personal/sensitive data of the trainees sent over a non-secure network;
- encryption of sensitive personal data transmitted between a remote server and a smartphone;
- enabling the mobile application operation offline. At the moment, mobile application is not able to operate without access to the Internet. This problem would be solved by introducing mechanisms for collecting data locally and synchronizing when the device is connected to the Internet;
- constant, cyclic expansion of sensors to be used and sensor signals to be recognized;
- introduction of *privacy policies* in the Web application so that the user decides what data may be available to his trainer;
- the ability to assign a user to a few trainers in a Web application.

At a further stage of the system development, with the efficient knowledge gained from data analysis, the role of the human trainer could be gradually reduced and replaced by AI solutions.

4 Potential MKUFUNZI Applications

Apart from utilizing the MKUFUNZI by individual users to measure and monitor vital parameters during physical exercises and to manually entering extra data into medical file, there are some other possible applications of the system.

The first and foremost is its application in fitness centers and gyms, where one trainer may control many trainees in the room. The trainer will be able to observe in real time the progress of several exercising people equipped with the training device and a mobile application.

The system could be used to mobilize children and pupils to exercise during physical education lessons. Children love new technologies and will probably be

fascinated by the possibility of utilizing the training device and smartphone during classes. The teacher will have a permanent look on children. Observing their activity in the Web application the teacher will be able to fairly evaluate the advancement of pupils and better motivate those avoiding exercises and to control active participation in classes.

MKUFUNZI can also serve seniors in hospitals, sanatoriums spa centers, providing staff access to electronic records of patients and allowing physiotherapists real time monitoring of their vital parameters and their pace of performed exercises. In this way the individual approach to a patient will be facilitated. The supervisor will be able to encourage some trainees to be more active or to stop others exercising too intensely.

The Assistive Integrated Personal Trainer System would also be very useful for seniors staying at home who, thanks to a mobile platform and a Web application could be under the constant remote control by medical or non-medical staff.

5 Conclusions

In the paper the prototype of the Assistive Integrated Personal Trainer System called MKUFUNZI has been presented. It is a feasibility study of personal informatics device assisting people of any age (also seniors) in their physical activities remotely supervised by the professional trainer and enabling an insight, via the Web, into their current and past results. The system is independent from commercial hardware and software and has got positive opinions from first users. By data gathering in a database a modeling the trainee activities is possible for better personalization and for system transition towards an intelligent computer assistant.

References

1. 7 minute workout. <https://itunes.apple.com/us/app/7-minute-workout/>. Accessed 9 Feb 2016
2. Apple.com, Apple introduction to health. <http://www.apple.com/ios/health/>. Accessed 9 Feb 2016
3. Cecchinato, M., Cox, A.L., Bird, J.: Smartwatches: the good, the bad, the ugly ?, CHI '15 Seoul, Extended Abstracts (2015)
4. Endomondo. <https://www.endomondo.com/>. Accessed 9 Feb 2016
5. Epstein, D.A., Ping, A., Fogarty, J., Munson, S.A.: A Lived Informatics Model of Personal Informatics, UbiComp '15. Osaka, Japan (2015)
6. King, K.: Physical activity for an aging population. *Public Health Rev.* **32**(2), 401–426 (2010)
7. Li, I.: Personal informatics and context: using context to reveal factors that affect behavior. Ph.D. Thesis, Carnegie Mellon University, Pittsburgh (2011)
8. Li, I., Dey, A., Forlizzi, J.: A Stage-based model of personal informatics systems, CHI 2010, Atlanta, pp. 557–566. Georgia, USA (2010)
9. Material Design. <http://developer.android.com/design/material/index.html>. Accessed 9 Feb 2016

10. MS HealthVault: <https://www.healthvault.com>. Accessed 9 Feb 2016
11. Ohlin, F., Olsson, C.M.: Intelligent computing in personal informatics: key design considerations. In: Proceedings of the 20th International Conference on Intelligent User Interfaces, IUI 2015, Atlanta, GA, USA (2015)
12. S-Health. <https://shealth.samsung.com/>. Accessed 9 Feb 2016
13. Starner, T.: How wearables worked their way into the mainstream. *IEEE Pervasive Comput.* **13**(4), 10–15 (2014)
14. Zdrowia Nieskończoność. <https://play.google.com/store/apps/details?id=com.droidinfinity.healthplus>. Accessed 9 Feb 2016

Quantitative Validation of Gait and Swing Angles Determination from Inertial Signals

Paula Stepien, Zuzanna Miodonska, Agnieszka Nawrat-Szoltysik,
Monika N. Bugdol, Michal Krecichwost, Pawel Badura,
Piotr Zarychta and Marcin Rudzki

Abstract The still increasing life length expectancy creates new challenges in the field of senior care. It encourages researchers to provide the nursing homes and senior care assistants with tools that will both, rise an alarm in case of a sudden fall and collect data for long-term diagnosis of the declining motor abilities like the number of steps taken per day or changes in some gait parameters. This paper presents a quantitative validation of a remote system for activity monitoring of the elderly based on inertial sensors. It focuses on features connected to walk quality such as number of steps and the swing angle outlined by an ankle in the sagittal plane during walk. A measurement protocol is proposed, a validation method is described and the obtained results are discussed.

Keywords Gait parameters · Inertial sensors · Activity monitoring · Signal processing

1 Introduction

Safety and well-being of the elderly stands for one of the fastest emerging issues in today's healthcare. Consistent progress in medicine and technology results in still increasing number of older people. This causes a need to develop medical and social branches offering a proper kind of care. Economical and—more importantly—psychological motives are the main foundations for the ambient assisted living (AAL) systems for telecare purposes [11]. They are supposed to provide tools for ensuring

P. Stepien (✉) · Z. Miodonska · M.N. Bugdol · M. Krecichwost ·
P. Badura · P. Zarychta · M. Rudzki
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: paula.stepien@polsl.pl

A. Nawrat-Szoltysik
The Jerzy Kukuczka Academy of Physical Education in Katowice,
Mikolowska 72b, Katowice, Poland

well-being of the elderly in their home environment. Several requirements are formulated for AAL architecture, e.g. sufficient level of simplicity, imperceptibility for the subject, low cost and energy constraints. Commercial systems offer safety functionalities, e.g. fall detection, automatic alerts, short-term threats detection, etc. [5, 15, 19]. Higher level of analysis employs data from multiple sources and is performed in order to draw conclusions about the patient health or its changes in various time periods [8, 17]. Multiple devices can be used as sensors distributed around the patient, e.g. cameras [6], passive infra-red movement detectors [5, 18] radio wave or ultrasound beacons [10]. However, thanks to miniaturization, the inertial sensors are probably the most exploited type of data sources [4, 7, 13, 14] also as a component of more complex systems [1].

Not only the residents of the assisted living facilities but also seniors under the supervision of family or friends exhibit interest in systems that might give them the sense of security without burdening their relatives. A tool that could both rise an alarm in a dangerous situation and inform in the case of declining motor abilities is being developed by researchers and commercial facilities worldwide.

To meet the expectations mentioned above, a system for the remote monitoring of geriatric patients in their home environment has been developed [12]. A physician operating on the hospital/telecare center side is provided with the data employing multiple sources to get a support in his/her diagnosis. The main source of information is a set of inertial sensors attached to a patient body, collecting accelerometric and gyroscopic signals and sending them to a processing unit. Various processing techniques extract information from the signals and provide indications, metrics and features interpretable in a different timespan: from short-term detections of incidents (such as faints, falls or slips), through middle-term notifications on patient hourly or daily activity, to long-term supervision. A long-term monitoring and inference system is supposed to employ both, the abstract quantitative measures extracted directly from the accelerations, angular velocities and their derivatives and more complex data calculated by dedicated algorithms, e.g. specific activity distribution [9] or automatic assessment of motion or balance abilities correlated with standard examination tools [2].

Measurements and primary or secondary features play an essential role at the basic level of an automatic assessment. A look at their suitability and validation stands for the main goal of this paper. As the most important metrics and features those related to gait are chosen, they include steps detection followed by angular and temporal gait features. A verification protocol is formulated in terms of necessary equipment, environment, and agenda. A group of seventeen geriatric patients was asked to participate in the experiment. The tests have been performed and the results statistically processed.

The paper is organized as follows. Section 2 presents the participants sample specification and briefly describes the data acquisition system. The data processing algorithms designed for the selected metrics and feature extraction, followed by description of the verification protocol are also introduced. Obtained results and statistics are shown in Sect. 3. Their discussion along with conclusions and possible future plans summarize the paper in Sect. 4.

2 Materials and Methods

2.1 Patients

The study group consists of 17 elderly residents of a Nursing Home (NH), Poland, at the mean age of 79 and the average of 3.4 years spent in the facility. Both, their fitness and the walk quality are evaluated by an expert. Table 1 presents the patients and their condition (two rightmost columns), where value 1 stands for very poor and 5 for very good.

2.2 Motion Data Acquisition

Patient motion is captured using a custom-made wireless data acquisition system consisting of five inertial sensors (3-axis accelerometer and gyroscope) and a control unit (Fig. 1a) [12]. Placement of the sensors on patient body together with the coordinate system used is shown in Fig. 1b. Such a placement allows not only the steps to be detected, yet also several gait features to be measured [3, 9]. Inertial signals

Table 1 Description of the study group

ID	Gender	Age	Years spent in NH	Fitness	Walk quality
NH01	M	70	2	2	3
NH02	F	67	2	2	3.5
NH03	F	70	1	1	3.5
NH04	F	70	1	2	5
NH05	F	86	16	2	3
NH06	F	74	1	3	2.5
NH07	F	91	5	3	1
NH08	M	78	4	2	2
NH09	F	70	2	2	4
NH10	F	77	1	2	2.5
NH11	M	86	1	1	4
NH12	F	87	2	1	4
NH13	F	92	6	2	3
NH14	F	85	4	2	3
NH15	F	85	3	1	4
NH16	F	75	5	1	3.5
NH17	F	81	2	1	5
Mean		79.06	3.41	1.76	3.32

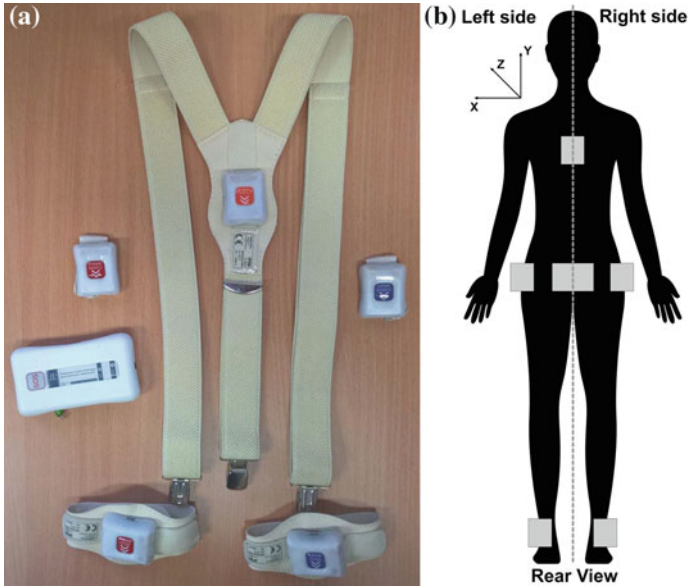


Fig. 1 Experimental setup. **a** Inertial sensor. **b** Placement of sensors

recorded by all sensors are stored in the internal memory of the control unit. After each measurement the collected data are transferred to a PC workstation for further analysis.

2.3 Motion Data Analysis

For the purpose of the experiment the following gait and fitness metrics are extracted from the acquired signals:

- the number of steps during an uninterrupted gait segment,
- the angle outlined by an ankle in the sagittal plane (the change of pitch $\Delta\theta_x$) during a step forward (for the right and left foot).

Preprocessing The absolute angles of sensors rotations are determined using a complementary filter, combining the low-pass filtered acceleration signal and high-pass filtered gyroscopic data [16]:

$$\theta = \alpha(\theta_{k-1} + \omega_k \Delta t) + (1 - \alpha)a_k, \quad (1)$$

where: ω_k is the angular velocity and a_k is the angle calculated using the acceleration, both in a specific plane and direction; $\alpha = T/(T + \Delta t)$, T is a time constant

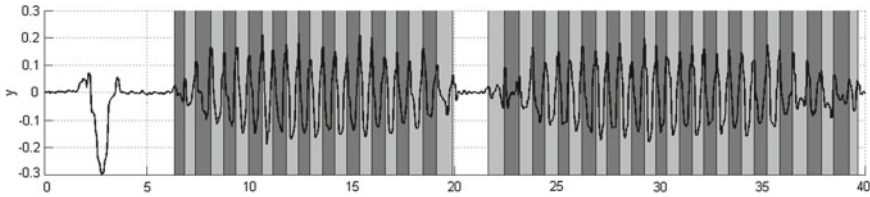


Fig. 2 Sample gait and steps detection results. Horizontal axis represents time in seconds, vertical represents y component of the acceleration signal from sensor on patient back. Grey segments represent gait, dark and light color distinguish steps done by left and right foot

(0.5 s here), and $\Delta t = 1/f_s$ is a time step. Three angles of rotation (roll θ_z , pitch θ_x , and yaw θ_y) are calculated using corresponding components of angular velocity and acceleration.

Step and Gait Detection The step detection algorithm operates on the y acceleration component measured by the sensor located on the back (Fig. 1b) [3]. Selection of this sensor for a rough step extraction is induced by its reduced sensitivity to unwanted limb movements, as well as consistent reaction to both feet strides. Local maxima are investigated in terms of their height and width (Fig. 2). A stride-related peak should be rapid: higher than surrounding signal fluctuations and relatively short. These assumptions eliminate gentle movements of the patient torso. If a peak meets stride assumptions, the adequate timestamp is assigned. A sequence of candidate steps is analyzed for inappropriate relations in time domain. Possible aggregations of close peaks are separated and replaced by single indications.

The resulting set of timestamps is considered definitive in terms of separate step detection and counting. However, a gait segments extraction continues for specific gait metrics extraction. This task is realized as a part of the activity recognition algorithm [9]. It groups the signal into five clusters: gait, postural transition (sitting to standing and standing to sitting), bow, other dynamic activities, and the lack of activity. The gait segment is found based on relationships between detected strides. As long as the neighboring strides' intervals are smaller than a constant value (set to 1.5 s), they are covered by a single walking phase. Note, that only at least 10-step segments are considered as walking and subjected to a fine analysis.

Swing Angles Determination The pitch is measured as a deviation from the vertical (Fig. 3). Therefore, the process of moving the foot forward after lifting it up is related to the decreasing pitch angle value. The calculation of the change of pitch is performed for individual gait cycles based on the angle signal yielded by the complementary filter. Characteristic points of steps—minima and maxima of this signal—are calculated, taking into account the average step duration specific for each patient and extracted using previously detected gait segments. The differences between adjacent minima and maxima are considered the angles outlined during the consecutive steps ($\Delta\theta_x$).

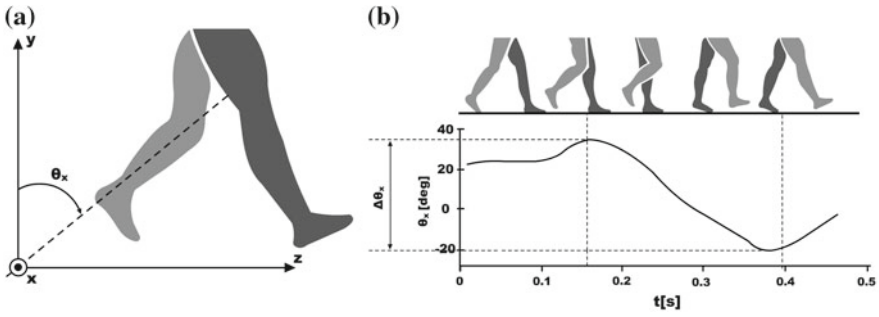


Fig. 3 The illustration of pitch θ_x at a specific moment (a) and $\Delta\theta_x$ is an angle outlined over time—during a step (b)



Fig. 4 Measurement track of the *synthetic* part

2.4 Measurement Track and Protocol

Two data acquisition protocols have been defined. The first one precisely defines the activities and is referred to as *synthetic*. The other one, *natural*, encourages the participant to perform daily activities. Each measurement is recorded using a HD camera. The recordings are presented to experts that evaluate both, the number of steps and the swing angles.

Synthetic The protocol contains several elements or situations that are typically present in the everyday life of an elderly person (Fig. 4) and includes:

1. The straight walk. On the floor two straight lines are placed in parallel and about 2 m apart that define a path where the examined person performs gait at different speeds. The path is placed perpendicularly to the camera axis, so that the patient movements are recorded from one side at a time. Such camera orientation makes it

possible to observe the gait characteristics and its symmetry. As the patient walks the path several times, the way of turning around at the end of the path is also observable. This path characteristics are very close to the natural motion paths traveled by the elderly in their rooms or in the hallways of the nursing facility.

2. The exercise chair, on which the examined person sits and performs basic exercises of the lower limbs. The tasks performed on the chair shows, how a person copes with changing position from standing to sitting and vice versa.
3. The couch, on which the examined person is lying on the back, then turns to one side of the body and then returns to sitting. This element of the track is to determine, how an elderly person can cope with changing their position while lying and with getting orthostatic.
4. The station consisting of a chair and the upper limbs rotor. This is one element of the set of exercises practiced by the elderly people living in the nursing home—it improves the joints mobility, blood circulation and tissue nutritioning. The part of the experiment concerning this station is performed in order to evaluate how such specific activities are interpreted by the algorithms under evaluation.
5. The stairs, that consist of 2 steps (seen in the background in Fig. 4). Despite a lift is available in the nursing home, walking a small number of stairs is still a part of everyday activities. However, not all patients can perform this task.

Natural This part resembles typical activities performed daily in the nursing home and consists of:

- putting on the jacket or a jumper,
- walking in the garden or within a nursery home building,
- walking upstairs,
- taking the elevator down.

The patients walk without assistance inside and outside the nursing home building. Three experts follow them taking notes and record the feet of the subject from behind. Each test is recorded with the main focus on the lower limbs movement (Fig. 5). The patients wear their private footwear not to influence their movements. The voice recording allows also the specific sound of the footsteps to be heard. Depending on the weather and the patient fitness, in some cases the *natural* phase of the experiment or some of its elements have to be omitted.

2.5 Data Verification

Steps The steps verification is performed on the data obtained during the *natural* phase of the protocol. The number of steps is counted by an expert based on the A/V recording and compared with the notes made during each measurement. Two modes of calculation are applied: a total number of steps taken during the entire examination



Fig. 5 Recording taken from the *natural* part

and the number of steps within extracted gait segments. The obtained number of steps is used as a gold standard for comparing the results of inertial signal processing methods.

Angles The verification of the swing angle calculation correctness is based on the measurements taken during the *synthetic* phase of the experiment. For each patient, 2–4 gait segments are taken (1–2 when the patient is visible in the camera image from the right side, 1–2 when the patient is visible from the left side). The number of selected gait segments depends on the number of steps registered by the camera during one segment—at least 3 full steps on each side are verified for a patient and for some patients it is impossible to find them within one segment. Recording of the entire step with a clear visibility of its beginning and end is the only specific step selection criteria.

The estimation of the angle outlined by the ankle during an individual step is made based on 2 selected movie frames (the beginning and the end of the step). The angle between the tibia and the vertical is measured on both frames and summed up. The overall measurement error within one step is assumed to be 5° . Based on the time of registration, the analyzed steps are subsequently measured out in the algorithm's partial results. Next, the angle values calculated by the algorithm are compared with the video-based values. The analysis is made independently for the left and right leg.

3 Results and Discussion

3.1 Step Count Verification Results

The data acquired from 13 elderly patients has been evaluated and presented in Table 2. The percentage of correctly recognized steps counted in the whole protocol and in the extracted continuous walk sample reached 97.14 % and 98.64 %, respectively. It indicates that additional activities performed beside the gait slightly worsen the results. No correlation between the quality of recognition and the fitness or walk quality of the patients was found.

3.2 Angles Verification Results

The basic statistical measures calculated for the expert results and calculated angles for the right and the left leg, respectively are presented in Table 3.

Neither measured nor calculated angles were normally distributed, as was concluded from the Lilliefors test at the significance level $\alpha = 0.05$. Therefore, the variables were compared using the Wilcoxon paired difference test. The results (pW) are presented in Table 4. The analysis suggests that there is no statistically significant difference between the medians of the results of the proposed method and the actual data—the proposed solution does not show a tendency to understate or overstate the estimated value. In addition, the Spearman and Kendall correlation coefficients

Table 2 The step detection summary

ID	Walk quality	All steps (%)	Steps in gait (%)
NH01	3.5	97.84	98.52
NH03	3.5	96.68	100
NH04	5	95.97	98.45
NH06	2.5	97.42	99.27
NH09	4	97.22	99.58
NH10	2.5	100	99.56
NH11	4	99.35	100
NH12	4	97.67	91.48
NH13	3	95.15	98.18
NH14	3	93.81	99.66
NH15	4	98.52	99.56
NH16	3.5	96.97	100
NH17	5	96.26	98.06
Mean	3.65	97.14	98.64

Table 3 The summary of basic statistical measures calculated for expert results and calculated angles

<i>Right leg</i>				
	Measured angle (°)	Calculated angle (°)	Absolute error (°)	Relative error (%)
Maximum	70.00	75.30	9.35	17.00
Minimum	30.00	30.80	0.04	0.06
Mean	53.83	53.89	1.49	2.91
Median	55.00	54.34	0.98	1.88
Standard dev.	10.15	10.26	1.48	2.90
<i>Left leg</i>				
Maximum	60.00	63.94	4.62	15.48
Minimum	25.00	22.00	0.12	0.31
Mean	51.88	51.78	1.65	3.35
Median	55.00	54.83	1.25	2.65
Standard dev.	9.66	9.65	1.14	2.75

Table 4 The results of the Wilcoxon test and Spearman and Kendall correlation coefficients for the measured and the calculated angles

		Right leg	Left leg
Wilcoxon test	pW	0.984	0.805
Spearman correlation	ρ	0.956	0.939
	p	6.568×10^{-35}	1.168×10^{-30}
Kendall correlation	τ	0.879	0.845
	p	8.779×10^{-22}	1.192×10^{-19}

were calculated. Based on these results (Table 4) there was no reason to reject the hypothesis of a lack of statistically significant difference between the calculated and the measured angles.

4 Discussion and Conclusions

The proposed quantitative validation method of the remote system for activity monitoring based on inertial sensors enables to validate the data acquisition and processing system utilizing the walk performance and its quality. The achieved results of 98.64 % of correctly counted steps encourage the further work. Additionally, no statistically significant difference between the medians of the results of swing angles determination algorithm and the actual data has been observed.

The presented measurement track permits the algorithms to be evaluated with high precision, however at a cost of expert-based time-consuming analysis of video recordings. Several other approaches are possible, yet the drawbacks are more evident. The most effective method to compare results is to use a specialized system for motion tracking that uses reflective markers. A set of infra-red cameras allows the gait to be analyzed. However, such markers have to be placed directly on the patient's skin and be visible to the cameras. Using adhesives to bond a marker to skin of an elderly person is troublesome mainly due to the poor condition of his/her skin. Because the safety is of the highest concern such method is to be disregarded. Additionally, the requirement that markers have to be visible does not allow patients to wear their causal clothing, thus making them more reluctant to participate in the experiment. Finally, such a system requires lengthy installation and calibration.

Another approach would be to use solutions for physical activity monitoring already present on the market. Yet simple pedometers occur to be prone to errors resulting from the specific type of steps made by elderly people. They are unable to recognize steps made by a person that lumbers or shuffles. Therefore the presented approach fulfills its tasks despite some inconveniences and allows the measurement and data analysis methods to be verified.

Further research involves an attempt to obtain motion data using standard posture stability tests (Berg and Tinetti) for the same patient group. The goal is to make the evaluation results independent from the potential overrating or underrating of the tasks by a physician.

As a final remark, an interesting observation was made by the assistants in the nursing home. The residents reported the growth of their sense of security and the feeling of being important during this study, which is an additional psychological benefit of being in the focus of younger scientists.

Acknowledgments Project co-financed by the European Regional Development Fund under the Operational Programme Innovative Economy, project no. POIG.01.03.01-24-061/12. The authors wish to thank the medical staff of the Nursing Home Święta Elżbieta in Ruda Śląska for the possibility of conducting the experiments.

References

1. Augustyniak, P., Smolen, M., Mikrut, Z., Kantoch, E.: Seamless tracing of human behavior using complementary wearable and house-embedded sensors. *Sensors* **14**(5), 7831–7856 (2014). <http://www.mdpi.com/1424-8220/14/5/7831>
2. Badura, P.: Accelerometric signals in automatic balance assessment. *Comput. Med. Imaging Graph.* **46**, 169–177 (2015). <http://www.sciencedirect.com/science/article/pii/S0895611115000932>
3. Badura, P., Pietka, E., Franiel, S.: Acceleration trajectory analysis in remote gait monitoring. In: *Engineering in Medicine and Biology Society (EMBC) 2014 36th Annual International Conference of the IEEE*. pp. 4615–4618 (2014)

4. Bussmann, J., Martens, W., Tulen, J., Schasfoort, F., van den Berg-Emons, H., Stam, H.: Measuring daily behavior using ambulatory accelerometry: the activity monitor. *Behav. Res. Methods Instrum. Comput.* **33**(3), 349–356 (2001)
5. Charlon, Y., Fourty, N., Bourennane, W., Campo, E.: Design and evaluation of a device worn for fall detection and localization: application for the continuous monitoring of risks incurred by dependents in an alzheimer's care unit. *Expert Syst. Appl.* **40**(18), 7316–7330 (2013)
6. Costa, Á., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. *Expert Syst. Appl.* **39**(15), 12192–12204 (2012)
7. Godfrey, A., Conway, R., Meagher, D., O'Laighin, G.: Direct measurement of human movement by accelerometry. *Med. Eng. Phys.* **30**(10), 1364–1386 (2008)
8. Lee, S.W., Song, S.H.: A monitoring system for assessing life pattern of the elderly living alone. *Adv. Inf. Sci. Serv. Sci.* **3**(7), 311–323 (2011)
9. Lubina, P., Rudzki, M.: Artificial neural networks in accelerometer-based human activity recognition. In: *Mixed Design of Integrated Circuits Systems (MIXDES)*, 2015 22nd International Conference. pp. 63–68 (2015)
10. Marco, A., Casas, R., Falco, J., Gracia, H., Artigas, J.I., Roy, A.: Location-based services for elderly and disabled people. *Comput. Commun.* **31**(6), 1055–1066 (2008). <http://dx.doi.org/10.1016/j.comcom.2007.12.031>
11. Mitas, A.W., Rudzki, M., Skotnicka, M., Lubina, P.: Activity monitoring of the elderly for telecare systems—review. In: *Information Technologies in Biomedicine, Volume 4, Advances in Intelligent Systems and Computing*, vol. 284, pp. 125–138. Springer (2014)
12. Mitas, A.W., Rudzki, M., Wiclawek, W., Zarychta, P., Piwowarski, S.: Wearable system for activity monitoring of the elderly. In: *Information Technologies in Biomedicine, Volume 4, Advances in Intelligent Systems and Computing*, vol. 284, pp. 147–160. Springer (2014)
13. Oshima, Y., Kawaguchi, K., Tanaka, S., Ohkawara, K., Hikiyama, Y., Ishikawa-Takata, K., Tabata, I.: Classifying household and locomotive activities using a triaxial accelerometer. *Gait Posture* **31**(3), 370–374 (2010)
14. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: *In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence (IAAI)*, pp. 1541–1546. AAAI Press (2005)
15. Tasoulis, S., Doukas, C., Plagianakos, V., Maglogiannis, I.: Statistical data mining of streaming motion data for activity and fall recognition in assistive environments. *Neurocomputing* **107**, 87–96 (2013). *Timely Neural Networks Applications in Engineering Selected Papers from the 12th EANN International Conference* (2011)
16. Tseng, S., Li, W.L., Sheng, C.Y., Hsu, J.W., Chen, C.S.: Motion and attitude estimation using inertial measurements with complementary filter. In: *2011 8th Asian Control Conference (ASCC)*, pp. 863–868 (2011)
17. Virone, G., Alwan, M., Dalal, S., Kell, S.W., Turner, B., Stankovic, J.A., Felder, R.A.: Behavioral patterns of older adults in assisted living. *IEEE Trans. Inf. Technol. Biomed.* **12**(3), 387–398 (2008)
18. Yang, C.C., Hsu, Y.L.: Remote monitoring and assessment of daily activities in the home environment. *J. Clin. Gerontol. Geriatr.* **3**(3), 97–104 (2012). <http://www.sciencedirect.com/science/article/pii/S2210833512000494>
19. Yu, M., Rhuma, A., Naqvi, S., Wang, L., Chambers, J.: A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. *IEEE Trans. Inf. Technol. Biomed.* **16**(6), 1274–1286 (2012)

Application of Gait Index Assessment to Monitor the Treatment Progress in Patients with Cerebral Palsy

Katarzyna Nowakowska, Robert Michnik, Katarzyna Jochymczyk-Woźniak, Jacek Jurkojć, Marek Mandra and Ilona Kopyta

Abstract A three-dimensional analysis of movement is more and more often applied to assess locomotive functions of patients with cerebral palsy (CP). For the past few years there has been a growing interest in gait index methods which enable the evaluation of the patient's gait with the use of a single numerical value. This research aimed at the assessment of the gait of patients with CP at different stages of treatment on the basis of gait indices Gillette Gait Index (defined by Shutte) and Gait Deviation Index (defined by Schwartz and Rozumalski) as well as finding dependencies between the determined indices values. The studied group consisted of 56 healthy participants and 16 patients with cerebral palsy treated using botulinum toxin. On the grounds of the determined indexes, the improvement of gait was observed in the case of 75 % of patients with CP. Statistically significant differences were found between GGI and GDI values obtained in the initial investigation and after 6 months following the administration of botulinum (PRE and POST). There was a strong linear dependence observed between the GDI and GGI^* ($GGI^* = \ln(\sqrt{GGI})$) values. No significant dependence was observed between the percentage indices of improvement determined on the basis of the GGI and GDI results obtained (in the PRE and POST examination).

Keywords Gait analysis · GGI · GDI · Cerebral palsy

K. Nowakowska (✉) · R. Michnik · K. Jochymczyk-Woźniak · J. Jurkojć
Faculty of Biomedical Engineering, Department of Biomechatronics,
Silesian University of Technology, Roosevelta 40, Zabrze, Poland
e-mail: Katarzyna.Nowakowska@polsl.pl

M. Mandra
Department of Pediatric Neurosurgery, Medical University of Silesia,
Medyków 16, Katowice, Poland

I. Kopyta
Department of Paediatrics and Developmental Age Neurology,
Medical University of Silesia, Medyków 16, Katowice, Poland

1 Introduction

Cerebral palsy is a group of disorders of development and posture resulting from a non-progressive permanent damage to the brain in an early phase of development, which contributes to the limitations of the motoric functions. Cerebral palsy is one of the most common causes of motoric disability in the children's population (around 2 per 1000 live births). Apart from the motoric disfunctions, the children with CP often suffer from disorders of sensation, cognitive functions, reception of stimuli and seizure disorders. Cerebral palsy is an irreversible neurological disorder, however, a quick and accurate diagnosis as well as adequate treatment and rehabilitation enable the compensation of many functions and make it possible to avoid permanent damage to the motoric system [3]. For diagnosing and assessing the treatment progress of patients with cerebral palsy, observation methods are the ones most frequently used. The most popular methods qualifying patients in accordance with certain scales include: GMFCS (Gross Motor Function Classification Scale), GMFM (Gross Motor Function Measure), PEDI (Pediatric Evaluation of Disability Inventory), FAQ (Functional Activities Questionnaire) and Ashworth scale which is a measure of muscle spasticity. Using observational scales is, however, a very subjective method, whose result depends directly on the researcher.

The analysis of gait constitutes a very important element of the examination of patients with cerebral palsy. In order to conduct a quantitative assessment of locomotive functions more and more often the three-dimensional analysis of movement is used. Another very important aspect of the evaluation of locomotive functions relates to the determination of the loads carried by the muscular and skeletal system as well as forces generated by individual muscles [4, 6]. The quantitative analysis of gait provides information which makes it possible to define the level of pathology and its development in time. It also enables the assessment of the treatment progress, diagnostics support and rehabilitation planning [7, 13, 14]. Due to the fact that the gait analysis generates a large amount of data, which are difficult to be quickly analysed by doctors or physiotherapists, the necessity arose to build a simple quantitative measure which would describe a patient's gait in a global (overall) way [1]. Trying to meet the new requirements the researchers determined a few indices of locomotive functions which present the gait variables in the form of a single numerical value. The most often determined gait indices are as follows: Gillette Gait Index [10] and Gait Deviation Index [11]. The indices determined by the authors show to what extent the gait of a patient under examination diverges from an average regular gait. Input data in the process of creation of the locomotive functions indices include different sets of space-time parameters as well as kinematic and dynamic parameters. Gait Deviation Index is a new quantitative measure which by means of the recognition of standards compares 9 kinematic gait variables (describing the kinematics of the pelvis, hip joint, knee joint and tarsal joint) of a given patient to the set of variables obtained for a group of patients with regular gait. In order to determine an Gillette Gait Index one uses 4 space-time parameters and 12 kinematic parameters. The index is dedicated to the evaluation of the locomotive functions of

patients with cerebral palsy [10, 12]. Moreover, it was proven that it can be used in the clinical assessment of the gait of adults with e.g. brain damage and spinal cord injuries [2], or patients with neoplasms of the central nervous system [12]. It was found that there is a big dependence between the results of the Gillette Gait Index and Edinburgh Visual Gait Scale [2]. In addition to that, the tests of the patient's with CP gait performed by Molloy et al. showed a strong connection between the determined values of GDI index and motoric functions expressed by means of the following scales: GMFM i GMFCS [8]. The authors of the Gait Deviation Index, Schwartz and Rozumalski found that its values differ between subsequent adjacent levels of the FAQ scale (Gillette Functional Assessment Questionnaire), whereas Cimolin et al. in their work [1] examined the efficiency of the application of the Gait Deviation Index to quantitative assessment of changes occurring after a surgical operation of the elongation of the gastrocnemius muscle fascia in patients with CP. In order to evaluate the results of the gait of a particular population by means of indices methods, it is important to verify which locomotive functions index is the most sensitive. Moreover, it is essential to be able to properly interpret the information provided by a given gait index [5]. The indices do not reflect the influence of individual muscle physiology and anatomy. In the past few years many works have been published in connection with the application of gait indices to the functional assessment of the limitations of locomotive functions. However, there are very few works examining the dependence between individual gait indices. Moreover, it seems that there are no investigations in which gait index methods would be used to evaluate the progress of treatment with botulinum in patients with cerebral palsy.

Taking into consideration the above, within the framework of this work the authors of this paper decided what follows:

- to assess the gait of patients with palsy at different stages of treatment with botulinum toxin on the basis of the Gillette Gait Index and Gait Deviation Index;
- to find dependencies between the determined GGI and GDI values;
- to verify the possibility of using gait indices in the assessment of the progress of treatment with botulinum toxin in patients with CP.

2 Materials and Methods

The studied group was composed of 56 participants having regular gait and 16 patients with cerebral palsy qualified for the treatment using botulinum toxin. The characteristics of the group studied have been presented in Tables 1 and 2.

The tests of the locomotive functions were conducted at the Upper Silesian Children's Health Centre in Katowice with the use of BTS Smart-D optical system. This system is used for the analysis of movement and consists of: a set of markers placed in precisely defined anatomical points on the body of the examined person, a set of 8 cameras recording the changes of location of the markers (of the frequency of the recording sampling 250 Hz), two vision cameras as well as dynamometric platforms

Table 1 Characteristics of patients with CP and reference subjects

	Summary statistics	CP	Reference
Number (Men/woman)		16 (9/7)	56 (28/28)
Age (years)	Mean (sd)	5 (3)	11 (3)
BMI (kg/m ²)	Mean (sd)	16 (3)	19 (3)

Table 2 The characteristics of the patients with CP

	Summary statistics	Hemiplegia	Diplegia	Quadriplegia
Patients treated using botulinum toxin	Percent (number of patients)	31 (5)	56 (9)	13 (2)

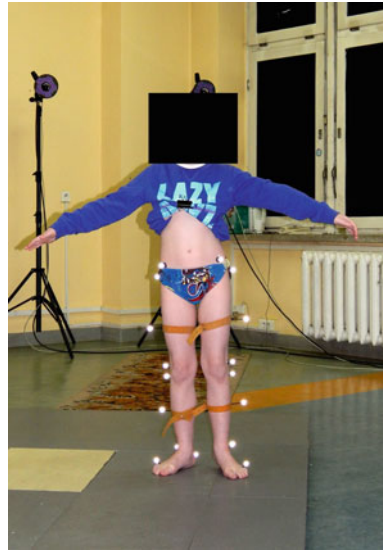
of the Kistler company. Prior to the gait examination, the measurements of the necessary anthropometric values were conducted. Next, according to Davis' protocol, passive markers were placed on the patient's body in precisely defined anatomical points. During the measurement the markers were recorded by BTS Smart system. During the test each patient walked along the measurement path at their natural speed. All participants walked without shoes and did not use any support devices. The illustrations below (Figs. 1 and 2) present a diagram of the measuring station and distribution of the markers on the patient's body.

Ethical approval for this study was obtained from the Ethics Committee of Medical University of Silesia (no. KNW/0022/KB1/19/11). The tests of the gait of patients with cerebral palsy were performed twice: before the beginning of the treatment

Fig. 1 Measuring station—BTS Smart optical system



Fig. 2 Distribution of markers on the patient's body according to Davis' model



(i.e. before the administration of botulinum toxin) and 6 months after the administration of botulinum toxin. Basic space-time and kinematic gait parameters were determined for all participants tested. Using author's software applications written in Matlab environment, gait indices (Gillette Gait Index and Gait Deviation Index) were determined for each tested person. On the basis of the GGI and GDI results obtained for the patients with CP, the percentage indices of the improvement between the examination conducted before (PRE) and after (POST) the botulinum treatment was determined according to the following equations:

$$I_{GGI} = \frac{GGI_{POST} - GGI_{PRE}}{GGI_{PRE}} \cdot 100\%, \quad (1)$$

$$I_{GDI} = \frac{GDI_{PRE} - GDI_{POST}}{GDI_{PRE}} \cdot 100\%, \quad (2)$$

The results obtained were subjected to the statistical analysis. Correlation (by Pearson) between the obtained index values (GDI and $GGI^* = \ln(\sqrt{GGI})$) and percentage indices of improvement were calculated. The statistical significance was determined to be at the level of $p < 0.05$.

3 Results

The obtained GGI and GDI results have been presented in Table 3. Mean values for the control group were as follows: GGI— 15.7 ± 5.68 , GDI— 99.23 ± 8 . Whereas for the group of patients with cerebral palsy, the following mean values of the indexes were noted in the initial examination: GGI— 286.25 ± 266.94 , GDI— 77.35 ± 8.55 (Table 3). The relatively large standard deviation GGI value results from the fact that the values in the group of patients were within the range of 57.64–1054.7. Such a large diversity of index results for a group of patients (treated with the same method) stresses the fact that cerebral palsy is a group of diverse symptoms and co-existing disorders with varying degree of intensity.

Figure 3 presents the values of GGI indices for patients with cerebral palsy (the results obtained prior to treatment) depending on a type of paresis.

Pearson's correlation coefficient was determined between the obtained results of the indexes GDI and GGI* ($GGI^* = \ln(\sqrt{GGI})$) for participants making up the reference group (the standard) and patients with cerebral palsy. Figures 4 and 5 show diagrams of dispersion of the results obtained. A red broken line in Fig. 4 marks mean values of the indices obtained for the control group. An average negative dependence was observed (classification according to J.Guilford'a) between the GDI and GGI* values of the group with regular gait ($r = -0.45$) and high correlations between the indexes obtained for the group of patients with CP: in the initial examination $r = -0.67$ and in the examination carried out 6 months after the administration of botulinum toxin $r = -0.64$. The determined correlation coefficients are statistically significant. On the basis of the obtained values of GGI and GDI indices, the percentage indices of gait improvement were determined for the studied group of 16 participants. On the grounds of the obtained improvement indices it was found that after 6 months after the administration of botulinum toxin 75 % of the patients improved their locomotive functions. The mean index of gait improvement for the

Table 3 GGI and GDI results

Gait index			Average	sd	Median	Max	Min
GGI	Patients with CP	PRE	286.25	266.94	199.31	1054.7	63.14
	Patients with CP	POST	264.07	322.07	127.22	1326.6	30.73
	Control group	–	15.7	5.68	15.15	30	7.46
GDI	Patients with CP	PRE	74.52	8.55	75.44	88.52	57.64
	Patients with CP	POST	77.35	11.58	78.17	96.3	52.84
	Control group	–	100	9.98	100.19	121.07	78.95

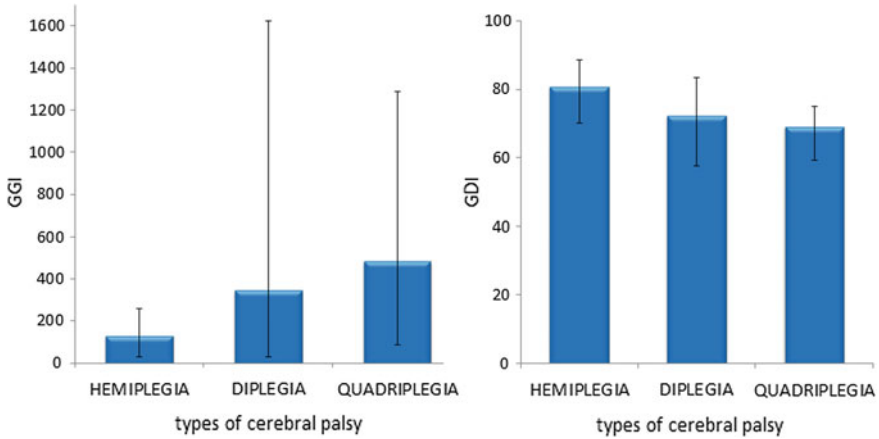


Fig. 3 Results of GGI and GDI for patients with CP depending on paresis type

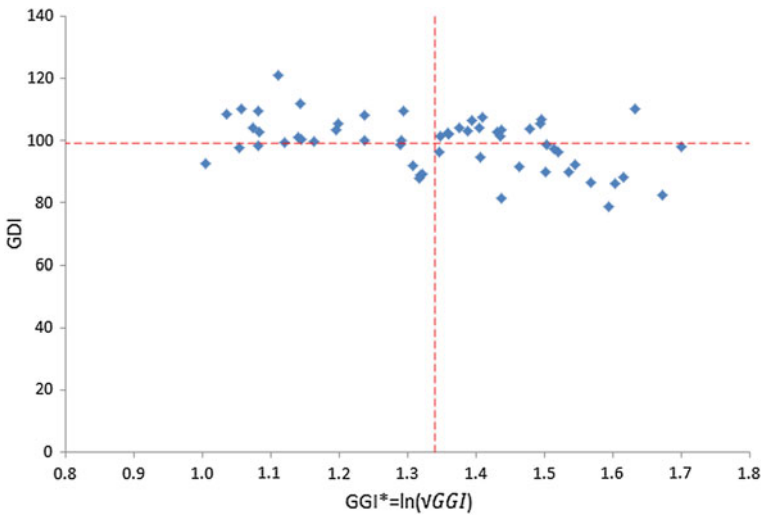


Fig. 4 Diagram of dispersion between GDI i GGI* values obtained for participants with regular gait

group of patients with cerebral palsy was determined on the basis of GGI index values and amounted to 17.63 % (28.09 % for the paretic limb or for the limb having a higher gait index in the initial examination, 7.17 % for the other lower limb), whereas GDI calculated on the basis of the obtained results was equal to 4.7 % (4.29 % for the paretic limb or for the limb having a lower gait index in the initial examination, 5.12 % for the other lower limb). Figure 6 presents the relationship between the percentage indices of gait improvement determined on the grounds of the GGI and GDI results obtained in the PRE and POST examination. A visible improvement of locomotion

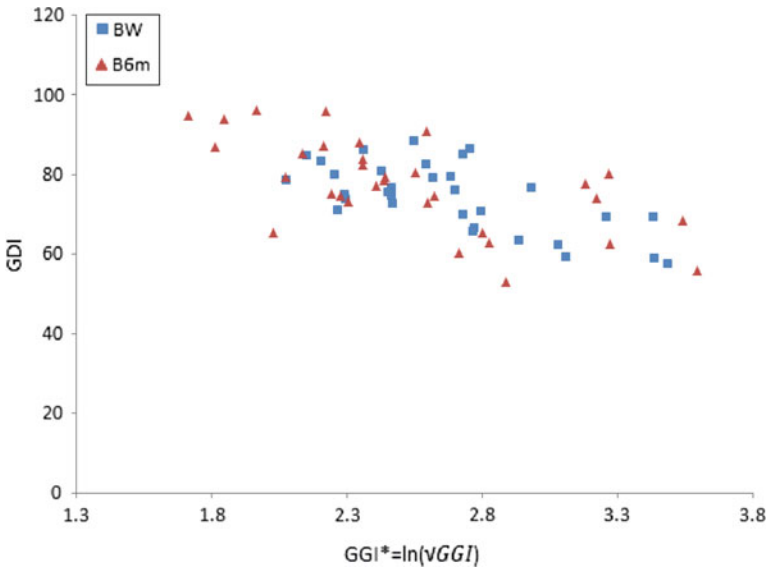


Fig. 5 Diagram of dispersion between GDI i GGI* values obtained for patients with cerebral palsy in examination BW and B6m

functions observed in 75 % patients can be regarded as a satisfactory result given the fact that botulinum toxin treatment aims to, first of all, decrease muscle spasticity and osteoarticular pains and, in the next phase, improve locomotion functions. On the basis of studies conducted as part of research for this paper, a study of locomotion functions was performed.

4 Discussion

Mean values of individual gait indices for the control group ($GGI=15.7 \pm 5.68$, $GDI=100 \pm 9.98$) coincide with values determined by the authors of other research papers including the authors who created the indices i.e. Schutte et al. [10] and Schwartz and Rozumalski [11]. GGI values determined for patients with CP in the initial examination (PRE) were within the range of: $63.14 \div 1054.7$, while in the tests conducted 6 months after the botulinum administration they equalled: $30.73 \div 1326.6$. Table 4 includes the comparison of the GGI values obtained for patients with CP with the results obtained by Schutte et al. and Romei et al. The values obtained for particular types of cerebral palsy are similar. In all research works the highest mean GGI value was obtained for patients with quadriplegia (tetraplegia), whereas the lowest deviation from regular correct gait was indicated in patients with hemiparesis [9, 10].

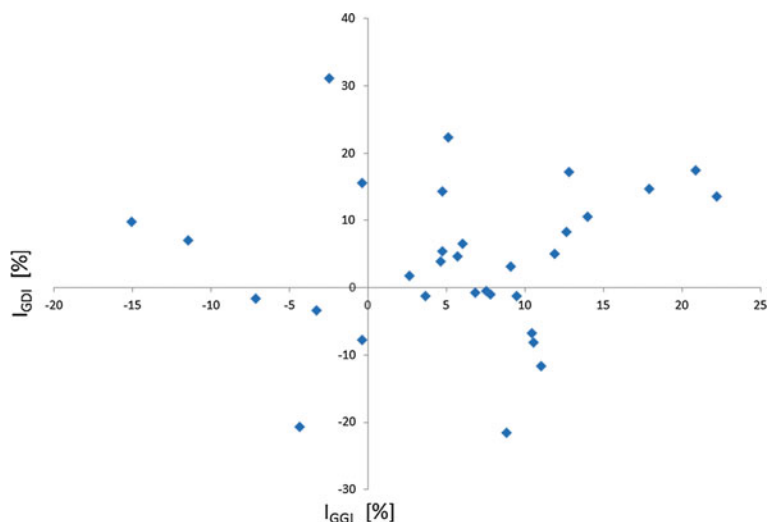


Fig. 6 Distribution of the percentage indices of gait improvement determined on the basis of GGI and GDI values for 16 patients with CP treated with botulinum toxin

Table 4 GGI values for patients with cerebral palsy depending on the type of paresis with reference to other authors' research

Paresis type	Schutte et al.	Romei et al.	Own research
Hemiplegia	166.7 (32.1 ÷ 1066.5)	189.28 (41.5 ÷ 435.5)	148.95 (63.14 ÷ 244.30)
Diplegia	279.4 (28.46 ÷ 1322.3)	278.12 (59.6 ÷ 789.5)	345.75 (81.82 ÷ 1054.7)
Quadriplegia	491.0 (121.5 ÷ 1195.0)	383.71 (177.4 ÷ 626.5)	361.75 (136.91 ÷ 674.12)

For patients with cerebral palsy the mean GDI values in the PRE and POST examination were as follows: 74.52 ± 8.55 i 77.35 ± 11.58 . It may be stated that before the administration of botulinum toxin the gait of patients with CP diverged from the mean of regular correct gait by $2.98 \times$ standard deviation, however, during the tests conducted 6 months after the injection of botulinum it diverged from the mean by $1.96 \times$ standard deviation.

The statistical analysis performed for the obtained results, i.e. a non-parametric Wilcoxon signed-rank test and a Student's t-test for dependent samples, showed statistically significant differences between the GGI and GDI values obtained in the initial examination and the tests carried out 6 months after the administration of botulinum toxin. Taking into account the above, it can be stated that the mean increase of GDI by 1 standard deviation is clinically significant. A statistically significant, moderately strong, negative correlation was observed between the GGI and GDI values. Taking into consideration the values of all the indices determined in this paper—it can be stated that a linear dependence $r^2 = 0.74$ was observed between GDI and GGI*. This value is slightly higher than the value obtained in the research done by Schwartz and Rozumalski [11], where $r^2 = 0.56$. The dependence between

the indices reveals that both measures serve the purpose of the assessment of gait pathology, i.e. deviations from the mean of regular correct gait. The range of the obtained results indicates that these indices reflect different aspects of gait pathology. The differences result from the fact that different input parameters are used to form such indices [11]. During the analysis of the indices values of gait improvement obtained on the basis of the GGI and GDI results of examinations PRE and POST, it was found that percentage values do not coincide strongly with one another. No statistically significant correlation was found between them. Taking into consideration the above, it was stated that the GGI and GDI enable the quantitative assessment of locomotive functions in comparison to the results of the control group. However, due to the differences in the obtained results of the treatment progress, it is advisable to check which of the indices constitutes the most sensitive measure for the evaluation of differences before and after the treatment of patients with cerebral palsy. In their work [5] McMulkin et al. made an attempt to compare GGI, GDI, GPS/GDI* values in the population of people with typical gait pathologies (7 groups of patients with different types of gait pathology who were subjected to surgical procedures) in order to determine their sensitivity and clinical significance. The authors found that the most sensitive measures in the evaluation of differences before and after the treatment are GPS and GDI/GDI*. The authors subjectively found that GDI/GDI* is the most intuitive measure of the evaluation of clinical changes, whereas Gillette Gait Index has non-parametric properties and its results are difficult to be intuitively interpreted [5]. It should be noted, however, that Gait Deviation Index has also some limitations in comparison with GGI index. This index shows to what extent the pathology has changed, however, it does not provide any information about the causes of such changes.

With relation to the issue of an unambiguous quantitative assessment of the progress of botulinum toxin treatment of patients with CP, it is planned in further research to carry out the evaluation of the tests' suitability for a particular test group to check which of the indices constitutes a better measure for the assessment of clinical changes.

5 Conclusions

The obtained values of GGI and GDI indices for the control group and for the group of patients with cerebral palsy are comparable to the results obtained by authors of other works [8–11]. For 75% of patients treated with botulinum toxin the improvement of locomotive functions was observed on the basis of the determined gait indices. The statistically significant differences were observed between GGI and GDI results obtained in the initial examination and the examination conducted 6 months after the administration of botulinum toxin (PRE and POST). There was quite a strong linear dependence $r^2 = 0.74$ observed between GDI and GGI* values. No significant dependence was observed between percentage indices of improvement determined on the basis of GGI and GDI results in PRE and POST examinations.

Acknowledgments The study was supported by the research grant DEC-2011/01/B/NZ7/02695 of the National Science Center.

References

1. Cimolin, V., Galli, M., Vimercati, S.L., Albertini, G.: Use of the gait deviation index for the assessment of gastrocnemius fascia lengthening in children with cerebral palsy. *Res. Dev. Disabil.* **32**, 377–381 (2011)
2. Cretual, A., et al.: Gillette gait index in adults. *Gait posture* **32**(3), 307–310 (2010)
3. Dudek, J., Chuchla, M., Snela., Szymczyk, D., Drużbicki, M.: Zaburzenia wzorca chodu u dzieci z mózgowym porażeniem (Disorders of gait pattern in children with cerebral palsy). *Przegląd Medyczny Uniwersytetu Rzeszowskiego*, no. 3, s. 317–322 (2009) (in Polish)
4. Jurkojć, J., Michnik, R., Pauk, J.: Identyfikation of muscle forces acting in lower limbs with the use of planar and spatial mathematical model. *J. Vibroeng.* **11**, 566–570 (2009)
5. McMullin, M.L., MacWilliams, B.A.: Application of the Gillette gait index, gait deviation index and gait profile score to multiple clinical pediatric populations. *Gait Posture* **41**(2), 608–612 (2015)
6. Michnik, R., Jurkojć, J., Pauk, J.: Identification of muscles forces during gait of children with foot disabilities. *Mechanika* **6**(80), 48–51 (2009)
7. Michnik, R., Nowakowska, K., Jurkojć, J., Jochymczyk-Woźniak, K., Kopyta, I., Mander, M.: Wykorzystanie metod modelowania obciążeń układu szkieletowo-mięśniowego u pacjenta z mózgowym porażeniem dziecięcym (The use of musculoskeletal system load modeling methods in patients with cerebral palsy). *Modelowanie Inżynierskie*, Tom 24, Zeszyt 55, s. 74–80 (2015) (in Polish)
8. Molloy, A., McDowell, B.C., Kerr, C., Cosgrove, A.P.: Further evidence of validity of the gait deviation index. *Gait Posture* **31**, 479–482 (2010)
9. Romei, R., et. al: Use of the normalcy index for the evaluation of gait pathology. *Gait Posture*, **19**, 85–90 (2004)
10. Schutte, L.M., Narayanan, U., Stout, J.L., Selber, P., Gage, J.R., Schwartz, M.H.: An index for quantifying deviations from normal gait. *Gait Posture* **11**, 25–31 (2000)
11. Schwartz, M., Rozumalski, A.: The gait deviation index: a new comprehensive index of gait pathology. *Gait Posture* **28**, 351–357 (2008)
12. Syczewska, M., Dembowska-Bagińska, B., Perek-Polnik, M., Kalinowska, M., Perek, D.: Gait pathology assessed with Gillette gait index in patients after CNS tumour treatment. *Gait Posture* **32**, 358–362 (2010)
13. Syczewska, M., Świącicka, A., Kalinowska, M., Gaff, K.: Zastosowanie ilościowej obiektywnej analizy chodu do oceny doboru zaopatrzenia ortopedycznego u dzieci z mózgowym porażeniem dziecięcym (analiza chodu w doborze zaopatrzenia medycznego w mpdz). (Use of quantitative, objective gait analysis for assessment of selection quality of AFOS in CP children (gait analysis in AFOS in CP)), *Fizjoterapia Polska*, vol. 4, no 4, s. 298–303 (2006) (in Polish)
14. Świącicka, A., Syczewska, M.: Assessment of the efficacy of rehabilitation with concomitant use of botulinum toxin in patients with cerebral palsy qualified for this therapy by means of objective gait analysis. *Med. Rehabil.* **11**, 9–16 (2007)

Recognising Gait Patterns of People in Risk of Falling with a Multi-layer Perceptron

Raoul Hoffmann, Christl Lauterbach, Axel Techmer, Jörg Conradt and Axel Steinhage

Abstract We present an approach to estimate a person's risk of falling by analysing gait recordings from a 6 m walk on a sensor floor. The risk of falls correlates with certain parameters of the human gait. Although these parameters are not measured directly with this sensor, their information is reflected in the data. The sensor floor works with a capacitance measurement principle and is sensitive to contact, such that persons standing, walking or lying on the floor are recognisable. In a preprocessing step, the person's position is determined from the sensor data and the spread of contact points calculated. This spread implicitly contains the step sizes for the step phase in which two contact points are present. For each experiment, the distribution of occurring spreads is binned and taken as an input feature vector to a feed-forward perceptron. The neural network was trained by backpropagation with gait recordings from persons in low risk of falling and persons in high risk of falling. In the dataset, subjects were labelled as being in high risk of falling based on the prevalence of diseases, falls that already happened, and expert opinions. Though in this setup the data was recorded within a controlled environment, the results are transferable to larger installations and long-term observation periods.

Keywords Gait analysis · Floor sensor · SensFloor · Perceptron

1 Introduction

For the elderly, falls are a serious health concern that often results in disabilities decreasing their quality of life [14]. Meanwhile, people wish to live an autonomous life in their homes as long as possible, and only consider to move into safer

R. Hoffmann (✉) · C. Lauterbach · A. Techmer · A. Steinhage
Future-Shape GmbH, Altlaufstr. 34, 85635 Höhenkirchen-siegersbrunn, Germany
e-mail: raoul.hoffmann@future-shape.com

J. Conradt
Neuroscientific System Theory, Technische Universität München, Karlstr. 45,
80333 München, Germany

© Springer International Publishing Switzerland 2016
E. Piętko et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_8

environments like senior residences as a last resort. The trade-off between self-determination and safety, fueled by ageing societies with a problematic ratio between carers and those in need of care resulted in research of technologies that support the elderly. A result of these developments is a large-area floor sensor, that can observe people's movements in their home and be used to assist them, by e.g. switching on lights at night to improve orientation, or automatically calling for help when a fall is detected. In this paper, we examine if the data delivered by this floor sensor is suitable for estimating the risk of falling. An autonomously performed, continuous unobtrusive assessment of the gait stability could deliver valuable hints to trigger early countermeasures to prevent falls, like intensified physiotherapy, technological assistance or to commission a carer.

2 Gait and Falls

Gait instability correlates with a tendency to trip and fall [16], affecting measurable characteristics of the gait cycle. Studies like [2, 3] found timing parameters to play an important role in the fall risk assessment. Variabilities in stride time (time between two subsequent initial contacts of the same foot), stance time (time between first and last contact of one foot and the floor), swing time and the ratio of stance to stride time were higher in persons who already experienced a fall. Furthermore, variabilities in step sizes seem to affect the gait stability [5]. Very low and very high variabilities of the step width were found to affect the fall risk in [1]. Other studies found the step length to influence the fall probability, as well as walking speed and cadence. Our hypothesis for this work is, that gait characteristics linked to fall tendencies influence the signal patterns generated by a person walking over a large-area sensor floor.

3 Data Acquisition

For acquiring the gait patterns, we used the floor sensor system SensFloor® [11] in a setup sized $6\text{ m} \times 1\text{ m}$, as shown in Fig. 2. SensFloor is a sensor that is capable of determining the positions of persons by performing electric capacitance measurements over large areas within a certain spatial resolution. The sensor system can be installed below common flooring and such be unobtrusively hidden in indoor environments. As shown in Fig. 1, SensFloor is a textile-based sensor with embedded electronic modules. The polyester fleece base is layered with a textile made electrically conductive by a metallic coating. Patterns are cut into the top layer that form the individual measurement areas and lines transmitting power to the electronic modules. The modules continuously measure the electric capacitance of the eight nearby triangular areas. Only when the capacitance value changes, a wireless message is sent out containing the module id and the capacitances of the fields. The result is an event-based behaviour of the sensor with a variable data rate dependent on the

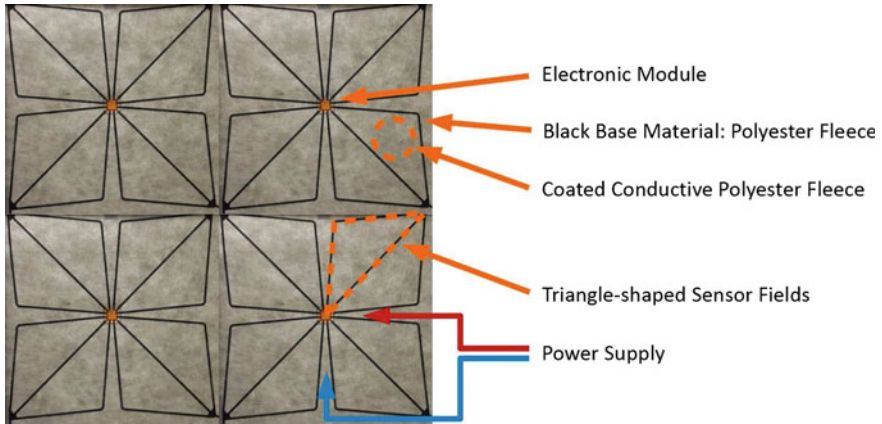


Fig. 1 SensFloor: the grey textile coating is cut to get traces for the power supply and the individual triangular sensing areas



Fig. 2 The gait recordings were taken with a SensFloor installation with a width of 1 and 6 m of length, covered by a blue carpet. The left photo was taken at the recording site in our offices, the one to the right in the senior residence

actual activity of people on the floor. Applications of SensFloor are fall detection and prevention in elderly care, person tracking [9] and robotic applications [10]. SensFloor was previously used to distinguish cats and persons moving on the floor [4]. Precise Human gait analysis with floor sensors can be achieved with higher resolution pressure-sensitive floor sensors [13]. In contrast to these systems, SensFloor features a lower resolution but larger area of measurement, and can be integrated in everyday living environments.

4 Experiment Setup

In cooperation with the senior residence *Wohnen am Schlossanger* in Höhenkirchen near Munich in Germany, 21 seniors participated in this study. At this senior residence, a fall statistic is archived for each individual inhabitant. The medium age of this group was 87 years at the time of recording, ranging from 74 to 95 years. According to the caring personnel and the ergo therapist, all participants are subject to a high risk of falling. The health records of the participants showed the prevalence of a possibly gait-influencing disease for each of them. The inhabitants of the residence who participated in the study suffered from Dementia, Parkinson's disease [8], osteoporosis, fractures and prostheses, or even combinations of them. From the group of 21, five persons already experienced a fall during their time in the senior residence. The group of persons with low risk of falling had a size of 10 participants with a medium age of 45 years.

For the recording of the gait patterns, the participants were asked to walk over the sensor area at their normal pace. Several of the participants from the seniors group relied on walking aids like walking canes or wheeled walkers. Some were accompanied by carers while walking over the floor to support them and prevent injuries in case of tripping or falling. Depending on their physique and endurance, every subject performed 2 or 4 walks over the area. The sensitive area was a SensFloor

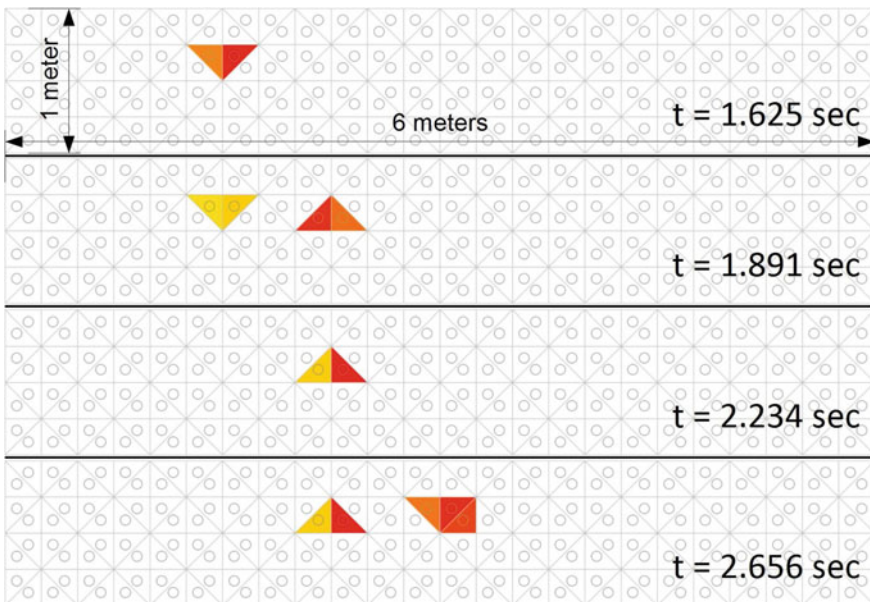


Fig. 3 Raw sensor data plotted for four consecutive moments of a walk over the area. *White* sensor fields are currently not occupied, colours in the range from *yellow* to *red* denote the capacitance level. The step phases *one foot on the ground* and *two feet on the ground* can be distinguished in this sample

installation with a length of 6 m and a width of 1 m, covered by carpet. A high resolution variant was used, such that the capacitance could be measured for 32 individual patches per square meter. The recordings of the senior group took place at the senior residence, the others at our offices. Figure 3 shows an exemplary time series of signals from the floor sensor. The recorded datasets were manually scanned for outliers in the sensor responses that can occur when sensor messages are not received. In total four experiments with such outliers were removed from the experiment sets. Furthermore, all experiments with rollators were removed, as the rollator wheels create a strong sensor response that interferes with the feature extraction step. After manual filtering, 91 experiments remained in the dataset pool, 45 from the group of fall-endangered seniors, and 46 from the group of persons not in danger of falling. The study is initial regarding the acquisition method with this type of large-area sensor. Comparable studies were performed with other kinds of floor sensors for differing applications [12].

5 Feature Extraction

By plotting the gait recordings as shown in Fig. 4, one can subjectively observe certain characteristic differences in the patterns generated by members of the two groups. In the group of non-fallers, distinct steps are clearly visible in the sensor data, which is mostly not the case for persons with a high risk of falling. In the latter group, persons typically had smaller step sizes and a lower lifting of the moving foot in steps, which results in a stronger sensor response in the swing phase. The feature extraction exploits the combined effect of foot lift height and step sizes on the sensor data to distinguish the two groups.

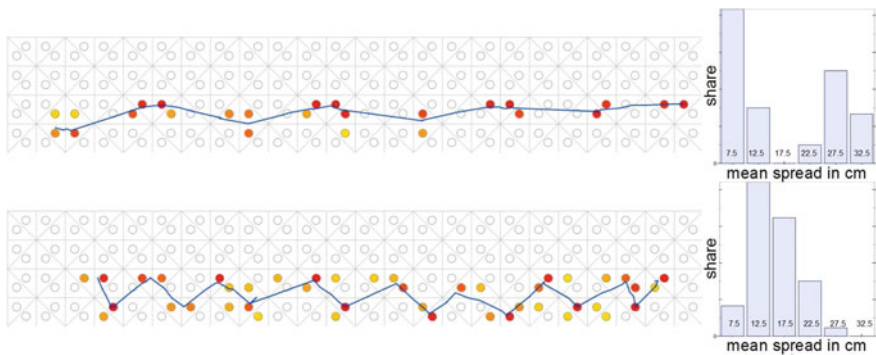


Fig. 4 These plots show tracks (blue line) and sensor activations (yellow to red circles) for two study participant. The top one is from the *low fall risk*-group, the bottom one from the *high fall risk*-group. Next to the tracking plots, the histogram shows the distribution of signal spreads that occurred

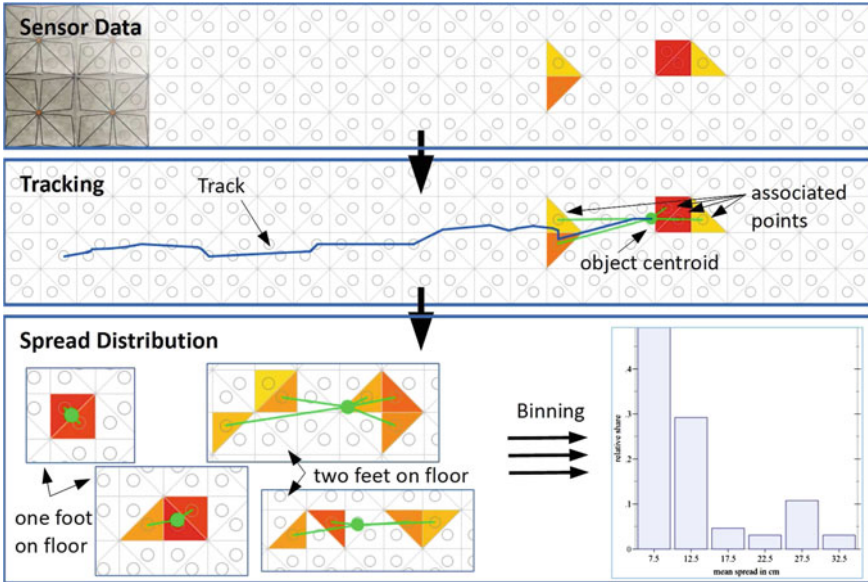


Fig. 5 Steps of the data processing: the sensor readings are recorded and processed offline to get the track of the object. For the track points over time, the mean distance of associated points to the centroid is calculated as a coefficient for the signal spread. The spread distribution is discretised in six bins and taken as input to a neural network

The feature extraction pipeline is a series of steps that calculates statistical information about the geometrical shapes of sensor activations from an experiment. At the end, a feature vector is obtained that contains discretised information about the distribution of spreads in the data. This feature vector serves as input to the multi-layer perceptron. The spread of a sensor reading at a time is defined as the mean value of all distances between a tracked object position and all associated points that contributed to the object position calculation (Fig. 5).

5.1 Preprocessing

The sensor data arrives in a format that allows for the assignment of a capacitance value to individual, triangular-shaped measurement areas. In a first step, these triangles are approximated and the centroid is calculated. All further calculations are based on the centroid positions with corresponding capacitance of the associated area. For an isosceles right triangle, the centroid is in $\frac{1}{3}w$ distance from the sides with w the length of the catheti of the triangle, as depicted in Fig. 6.

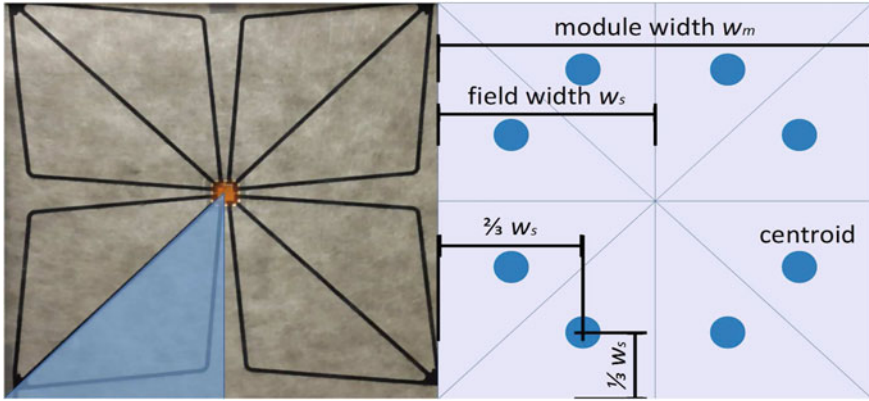


Fig. 6 The sensor field shape is approximated by isosceles right triangles with a cathetus length of half the module width. This way, the sensor field centroids are found at the drawn positions relative to the module edges

5.2 Tracking

The tracking of the person walking on the sensor area is achieved by calculating the center of sensor activations at the measurement points, weighted by their capacitance. This is comparable to calculating the center of gravity. Let

$$P_t = \{p_{i,t}\} = \{p_{0,t}, p_{1,t}, \dots, p_{n,t}\} \tag{1}$$

be the set of points

$$p_{i,t} = \{\mathbf{x}_i, c_{i,t}\} \tag{2}$$

at time t where $c_{i,t} \in [0, 128]$ is a measure of the current electric capacitance at the position $\mathbf{x}_i = \{x_i, y_i\}$ in 2-dimensional world coordinates with the origin in the corner of the measuring area. The tracked object position at time t is then:

$$\mathbf{x}_{obj,t} = \frac{1}{C_t} \sum_{i=1}^n c_{i,t} \mathbf{x}_i \quad \text{where} \quad C_t = \sum_{i=1}^n c_{i,t}. \tag{3}$$

C_t is the total capacitance of all measurement points at a time t .

5.3 Spread

For the subset $P_{a,t}$ containing all points in P_t with a capacitance $c_{i,t} > 0$, the euclidean distance to the tracked object's position is calculated. The spread s_t of sensor readings

at a time t is defined as the mean value of all distances between the object centroid $\mathbf{x}_{obj,t}$ and the associated points in $P_{a,t}$:

$$s_t = \frac{1}{|P_{a,t}|} \sum_{\{\mathbf{x}_i, c_{i,t}\} \in P_{a,t}} |\mathbf{x}_{obj,t} - \mathbf{x}_i|. \quad (4)$$

For every experiment and every recorded time with a change in the sensor signal, the spread was calculated this way. Finally, the spreads were discretised and summed up in 6 bins in the range from 5 to 35 cm, and normalised to get the discretised spread distribution. Figure 5 summarises the steps of the feature extraction process.

6 Classification Algorithm

The classification was implemented using the NeuroLab Python library [6]. From the artificial network types available in the library, a feedforward multilayer perceptron [7] was chosen and configured with 6 input neurons, 2 hidden layers with 7 neurons each, and a single output neuron. The transfer function for the neurons was set as hyperbolic tangent sigmoids. As input to the neural network the discretised spread distribution was taken. Every bin of the distribution corresponded to one input neuron. The distribution of spreads was normalised, such that the input range of the neurons is bounded to the range $[0 \dots 1]$. For the output neuron, a value of -1 was chosen to indicate a person endangered by falls, and $+1$ for a person with low risk of falling. The network architecture is shown in Fig. 7. This type of network was selected as it can also be used for regression tasks without much effort for reconfiguration. In future experiments, a larger data set could be learned by the neural network to not only output a fall risk class, but a continuous indicator of the fall risk.

7 Results

The neural network was trained with the dataset of ninety one individual experiments [15]. We applied a 9-fold cross-validation, such that the training and validation was performed nine times by splitting the dataset into 80 datapoints for training, and 10 datapoints for validation. This assured that every datapoint was exactly once part of a validation set. Taking the outcome of *person is in high risk of falling* as a positive result, the classification performed as follows: True Positive: 43, True Negative: 41, False Positive: 2, False Negative: 4. This calculates as a precision of 0.96 and a sensitivity of 0.91. The accuracy is 0.93. The output neuron's values were within a standard deviation of 0.05 around ± 0.99 .

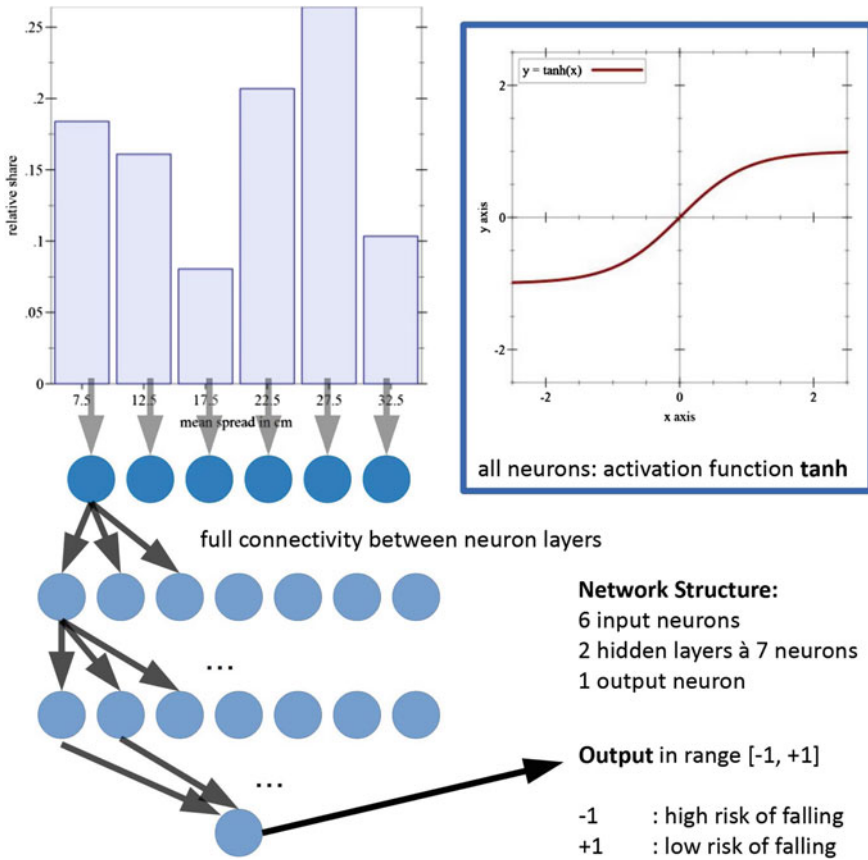


Fig. 7 Overview of the neural network structure and configuration. It is a multi-layer feedforward perceptron trained by backpropagation, implemented by the NeuroLab library and configured as shown here

8 Conclusion

We demonstrate that data from the floor sensor SensFloor is suitable to distinguish gait recordings from persons in high risk of falling from healthy subjects, though the spatial resolution of SensFloor is relatively low. A feature extractor is applied to the sensor data to extract the distribution of spreads of sensor activations. This spread distribution is discretised and taken as input to a neural network, which succeeds in performing the classification task with low errors. From the many gait properties that indicate a risk of falling, our feature extractor’s output reflects a combined measure regarding stance distances and foot lifting. Future work could address the extraction of timing-based properties from the sensor data, which also play an important role in gait analysis. Furthermore, a precise localisation of feet stances and follow-up

gait sequence recognition could extract some gait properties directly. Stroke patients often develop gait asymmetries, that could possibly be detected with a more detailed analysis. Currently, the fall risk assessment outputs a sharp binary distinction between *high risk of falling* and *low risk of falling*. By including persons with a history of more falls, it is thinkable to extend this approach to a regression task that outputs the risk of falling with finer granularity. Another application could be found in the evaluation of rehabilitation success, by taking multiple recordings during therapy periods. Although the gait recordings were taken in a controlled environment, it seems plausible to apply this method to long-term gait observations with arbitrary trajectories, for example in apartments equipped with SensFloor in larger areas. In these cases, changes occurring in the gait properties could be detected and e.g. be used to adjust the intensity of care.

Acknowledgments This work is supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) within the project *Cognitive VillAge: Adaptiv-lernende, technische Alltagsbegleiter im Alter* (CogAge, FKZ 16SV7311).

References

1. Brach, J.S., Berlin, J.E., VanSwearingen, J.M., Newman, A.B., Studenski, S.A.: Too much or too little step width variability is associated with a fall history in older persons who walk at or near normal gait speed. *J. Neuroeng. Rehabil.* **2**(1), 1 (2005)
2. Hausdorff, J.M., Edelberg, H.K., Mitchell, S.L., Goldberger, A.L., Wei, J.Y.: Increased gait unsteadiness in community-dwelling elderly fallers. *Arch. Phys. Med. Rehabil.* **78**(3), 278–283 (1997)
3. Hausdorff, J.M., Rios, D.A., Edelberg, H.K.: Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Arch. Phys. Med. Rehabil.* **82**(8), 1050–1056 (2001)
4. Hoffmann, R., Steinhage, A., Lauterbach, C.: Increasing the reliability of applications in AAL by distinguishing moving persons from pets by means of a sensor floor. In: *Proceedings SENSOR 2015, C5—Medical Systems*, pp. 436–440 (2015)
5. Lockhart, T.E., Woldstad, J.C., Smith, J.L.: Effects of age-related gait changes on the biomechanics of slips and falls. *Ergonomics* **46**(12), 1136–1160 (2003)
6. NeuroLab—a library of basic neural networks algorithms with flexible network configurations and learning algorithms for Python. <https://pythonhosted.org/neurolab/>. Accessed Jan 2015
7. Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**(6), 386 (1958)
8. Schaafsma, J.D., Giladi, N., Balash, Y., Bartels, A.L., Gurevich, T., Hausdorff, J.M.: Gait dynamics in Parkinson's disease: relationship to Parkinsonian features, falls and response to levodopa. *J. Neurol. Sci.* **212**(1), 47–53 (2003)
9. Sousa, M., Techmer, A., Steinhage, A., Lauterbach, C., Lukowicz, P.: Human tracking and identification using a sensitive floor and wearable accelerometers. In: *2013 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 166–171. IEEE (2013)
10. Steinhage, A., Lauterbach, C.: SensFloor® and NaviFloor®: robotics applications for a large-area sensor system. *Int. J. Intell. Mechatron. Robot. (IJIMR)* **3**(3), 43–59 (2013)
11. Steinhage, A., Lauterbach, C.: Sensfloor and navifloor: large area sensor systems beneath your feet. *Handb. Res. Ambient Intell. Smart Environ.: Trends Perspect.* **2**, 41–55 (2011)
12. Verghese, J., Holtzer, R., Lipton, R.B., Wang, C.: Quantitative gait markers and incident fall risk in older adults. *J. Gerontol. Ser. A: Biol. Sci. Med. Sci.* **64**(8), 896–901 (2009)

13. Webster, K.E., Wittwer, J.E., Feller, J.A.: Validity of the GAITRite® walkway system for the measurement of averaged and individual step parameters of gait. *Gait posture* **22**(4), 317–321 (2005)
14. Wild, D., Nayak, U.S., Isaacs, B.: How dangerous are falls in old people at home? *Br. Med. J. (Clin. Res. Ed.)* **282**(6260), 266–268 (1981)
15. Williams, D.R.G.H.R., Hinton, G.E.: Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986)
16. Wolfson, L.I., Whipple, R., Amerman, P., Kaplan, J., Kleinberg, A.: Gait and balance in the elderly. Two functional capacities that link sensory and motor ability to falls. *Clin. Geriatr. Med.* **1**(3), 649–659 (1985)

Determination of Loads in the Joints of the Upper Limb During Activities of Daily Living

Agata Guzik-Kopyto, Robert Michnik, Piotr Wodarski
and Iwona Chuchnowska

Abstract The elaborated methodology applied in the research integrates experimental measurements of kinematic parameters, which are then used in numerical computation. Measurements of kinematics of the upper limb motion were carried out during experimental studies using a MVN Biomech System. A mathematical model of movement of the upper limbs allowing the determination of force reaction in joints in a non-invasive way while performing selected routine activities is an important element of the applied research methodology. In the study the resultant reaction forces in the shoulder, elbow and wrist joints were determined and analyzed for the group of 8 healthy people as well as patterns of joint reaction force of the individual joints were developed.

Keywords Activities of daily living (ADL) · Upper extremity · Joint reaction forces

1 Introduction

Human movement is a multi-step process of activating a number of elements from the skeletal and muscular systems to an extensive system of nerves and the brain. An attempt to describe the movement of even one part of the body such as the upper limb, requires a complex kinematic and dynamic analysis [6, 20, 22]. The methodology which makes a kinematic analysis possible can be used due to optical or inertial systems for a motion analysis. Another type of describing human body movement and loads acting on musculoskeletal system are mathematical equations of motion [7, 8, 16, 21]. As a result, in a very simple way it is possible to calculate the trajectory of body movement, instantaneous acceleration and the velocity of the body. Robotics is based on these studies. According to a mathematical analysis, it is

A. Guzik-Kopyto (✉) · R. Michnik · P. Wodarski · I. Chuchnowska
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: Agata.Guzik-Kopyto@polsl.pl
URL: <https://www.polsl.pl/Wydzialy/RIB>

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_9

possible to create a machine that performs exactly the same movement a human at the particular moment [24]. The human upper limb is a mechanism which allows to perform complex manual activities. Human autonomy depends on its proper functioning. In their research Chan and Nicol [3] pointed out the hardships of motion analysis of the upper limb. In order to ensure a full recovery of patients after various types of surgery and diseases it is extremely important to position the upper limb. The first step to allow the analysis of an effective recovery was the introduction of kinematic sensors. This system turned out to be highly effective in the analysis of the movement of the upper limb and allowed to determine the correct movement patterns. In the studies by Pennestri and Stefanelli [24] a kinematic analysis has become a set of input data to a mathematical model, enhancing the credibility of the results. The model proposed by the authors was used to analyze driving in the driving simulator. The results of mathematical simulation were compared with the results obtained by using EMG. The conclusions in the study show the approximation of model calculation results to EMG with the use of kinematics research rather than classical mathematical method of determining motion. The next step in the methodology of model development was made by Albrecht and Arampatzis [1] who wanted to find out if general parameters of the person examined in a non-invasive way, could specify the size of the muscles of the particular limb in a given population. They conducted a study in a group of 13 men with the use of magnetic resonance imaging. During the research, the scientists have discovered that the size of the muscles of every human is an individual matter and these things cannot be considered as the average values in a group of patients. In the studies of the upper limb a research by Roman-Liu and Tokarski [27] was of great importance. They studied the influence of posture on the forces generated by the muscles of the upper limb. They carried out an analysis of the five movements of the upper limb: pushing force, lifting, and handle moments of pronation and supination. This kind of research has served to create optimal workplaces in which an employee can work efficiently using the least amount of energy. The development of movement patterns allows physiotherapists a better comparison of the results with the standard values and also enables a faster diagnosis [6, 20, 22]. Methodology assuming kinematic research which is an input data base for numerical models may prove to be extremely useful in assessing the degree of disability, or the evaluation of therapeutic effects [1, 7, 16].

The aim of the research was to develop patterns of joint reaction forces of the upper limbs while performing daily activities and to examine the loads difference in particular joints during performing of activities. The methodology of research and the results will be used in diagnosing patients with stroke and monitoring the progress of their recovery.

2 Materials and Methods

The research presented in this article was carried out in the Department of Biomechatronics of Silesian University of Technology and included a group of 8 men aged 22.7 (SD 2.32). In all subjects, there was no defect of the upper limb. In reference

Table 1 Description of subjects

	Mean	(sd)
Number (men)		8
Age (years)	22.3	(0.84)
Height (m)	1.80	(0.07)
Weight (kg)	72.87	(5.82)
BMI (kg/m ²)	22.40	(1.25)

to the study [2, 9, 17–19, 23] routine activities such as *drinking from a cup* and *lifting of an object* were selected. A group of subjects were asked to perform these two tasks. In both tests the subjects were seated on a chair, at a distance of 5 cm from the table. In the *drinking from a cup* test the task was to stretch out a hand, lift a cup of water, put it to the mouth, tip it and put it back, whereas the other hand did not perform any movement. In the second test, the subject lifted the object (a bottle filled with water), put it on an elevated surface, and then put it back (Fig. 1). The exact characteristics of the subjects are presented in Table 1.

Measurement of kinematics were carried out using MVN Biomech System consisting of 11 inertial sensors placed on the individual body segments. Each sensor was equipped with accelerometers, gyroscopes and magnetometers. Signals from the sensors are transmitted wirelessly via transmitters to a computer. Used system enables the measurement of the positions of anthropometric characteristic points of the upper limb joints and angles in the joints defined in accordance with the guidelines of the International Society of Biomechanics.

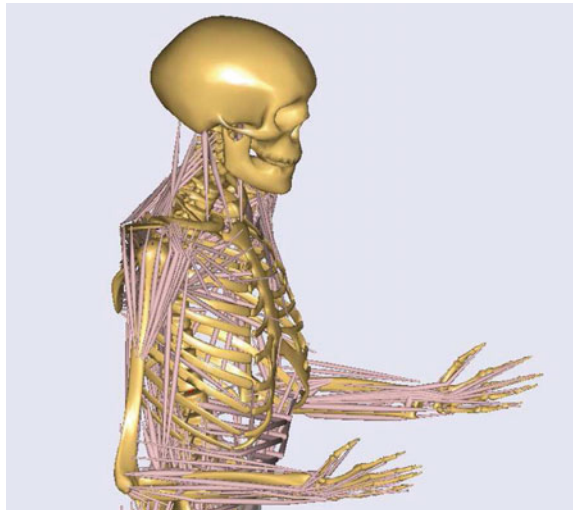
Measured and recorded kinematics was used to determine the input data and enable the simulation of loads of musculo-skeletal system of the upper limb. To carry out numerical calculations necessary to achieve the intended results a model in AnyBody Technology was developed. The structure of its record required to convert them to the

Fig. 1 Experimental tests for object lifting onto an elevated surface



forms generated from MVN Studio to a format readable by AnyBody programme. In numerical calculations Free Posture Move model supplied by AnyBody software was used. As in Quental [25] Pennestri [24] and Damsgaard [4] the model consists of segments (clavicle, scapula, humerus radius, ulna and hand) modelled in the form of rigid bodies connected by joints. The model is equipped with a module for adapting and scaling of the model segments dimensions for the anthropometric dimensions of subjects. The model consists of the following joints: sternoclavicular joint, shoulder joint, acromioclavicular joint, humero-ulnar joint, radio-ulnar proximal joint, wrist joint, radio-ulnar distal joint. All joints were modelled in the form of spherical and articulated connection forming biokinematic pairs of three (sternoclavicular articulation, and shoulder joints) and two (elbow and wrist joint) degrees of freedom. In the model, the following simplifications were applied: the symmetry in the construction of the limbs, two types of joint connections (spherical and articulated), the following forces were taken into consideration: gravity, inertia, and muscular forces and the forces caused of the articular surfaces. Frictional forces between individual elements (including friction forces in joints of the upper limb) were not taken into account, and the movement of soft tissue was omitted. Individual muscles were modeled by actons. The number of actons was different for the individual muscles. 46 major muscles of the upper limb divided into 144 actons were attached to the segments. The view of the model with the pictures of the segments for the sample person is shown in Fig. 2. The spatial arrangement of muscle structure and the position of each segment, the strength of the muscle and the location of muscles were developed largely on the basis of the Dutch Shoulder Group including research MAYO and VU [13] as well as literature posted on the AnyBody website [10].

Fig. 2 Model of the upper limb in AnyBody



The used model was a basis for the calculation of the inverse dynamic problem and kinematics of the upper limb while performing the registered activities. Thanks to the implemented programme AnyBody algorithms it is possible to calculate the loads the musculoskeletal system of the upper limb. Identification of muscular forces was carried out iteratively in each time step using inverse dynamics problem and static optimization. As a criterion for motion control objective function was assumed which minimize squares of muscle forces. Calculations with the use of AnyBody allowed to determine the forces generated in the individual muscles and reactions in the joints of the upper limb. Directions of reaction components in the joints have been defined according to the orientation of anatomical directions: proximal-distal, posterior-anterior and lateral-medial. Model calculations were performed for the whole group of subjects and for the analyzed activity in order to examine the differences between them. The result of the calculation are the values of reactions in the upper limb joints (shoulder, elbow and wrist). The calculation model takes into account the weight of the cup at 3 N force acting vertically down of the middle point of relief in one's hand in the direction of the force of gravity and the weight of the object (3 N). Analyses were carried out by comparing the shape of the courses of joint reaction forces of upper limbs which were obtained for individual and determining the maximal values of reactions in the joints.

3 Results

The reaction of forces in the joints of the dominant upper limb (in every person the right limb) are shown in the graphs (Figs. 3 and 4).

Both qualitative and quantitative assessment shows that for most subjects reaction forces in the individual joints are characterized by a high degree of similarity in relation to the others. It was observed that if the course of the reaction force of one of the subjects differed significantly from the other subjects, a similar discrepancy was also present in other joints. For example, for *drinking from a cup* (CD) the worst match course relative to the courses averaged over all the analyzed courses of the joints, characterized by forces of reaction were subjects b3 (green color), and b6. *lifting of an object* (OL) the worst match occurs in subject b6 (Figs. 3 and 4). The best match of courses of the reaction forces in all the joints in relation to the average values in the case of the *drinking from a cup* activity was b4, and for *lifting of an object* activity—b3 (see Figs. 3 and 4).

To better illustrate the proportion of the distribution of forces of reaction in different joints below are shown the courses of reaction forces generated by b4 and b3 while performing *drinking from a cup* and *lifting of an object* activities (Fig. 5).

Further analysis of the data was based on the average value of the maximal force reactions in the joints: the shoulder, elbow and wrist for the upper limb for each of the activities performed (Fig. 6).

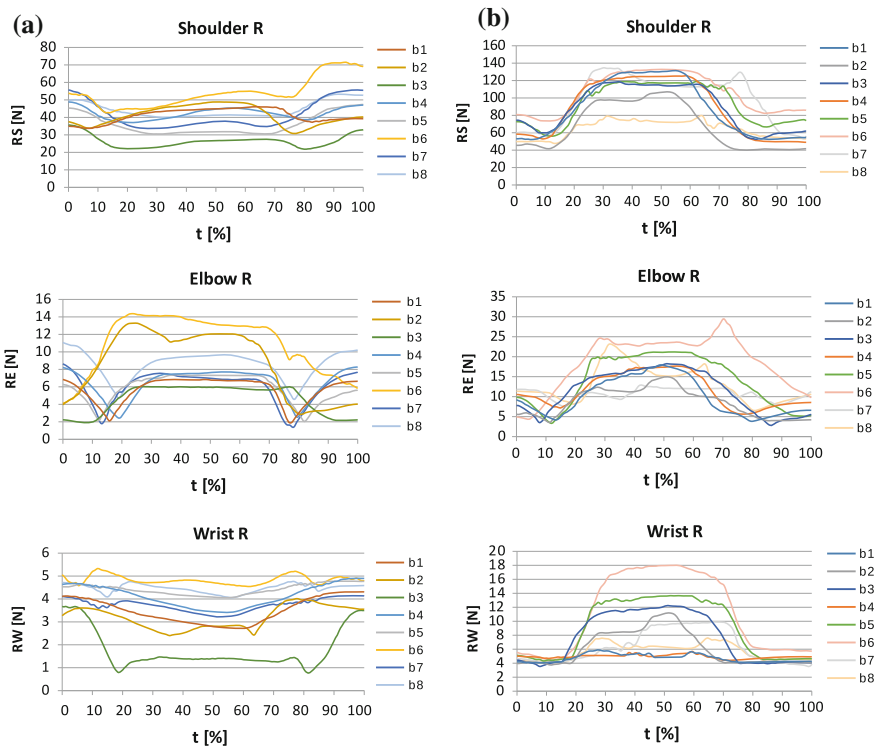


Fig. 3 Courses of reaction forces in joints: **a** shoulder, elbow and wrist during *drinking from a cup CD*, **b** shoulder, elbow and wrist during *lifting of an object OL*

Then the results were related to body weight of individual subjects. The mean value of the maximal reaction force of the joints as well as standard deviation were determined.

4 Discussion

Current knowledge which enables to determine exact loads in a muscular-skeletal system is still insufficient to clearly establish which muscle is responsible for performing a particular movement and what is a temporary value of its forces [5, 6, 11, 15, 25]. The muscle models developed in an approximate way allow for the simulation of the muscular-skeletal system [5, 11, 14, 19]. The calculation of the inverse dynamics problem is a highly complex issue which requires high computational power. Measured and recorded kinematics in this work was used to determine the input data and enable the simulation of the actual upper limb motion. Kinematic data were also used by Holzbaur and Murray [8]. In their research special attention

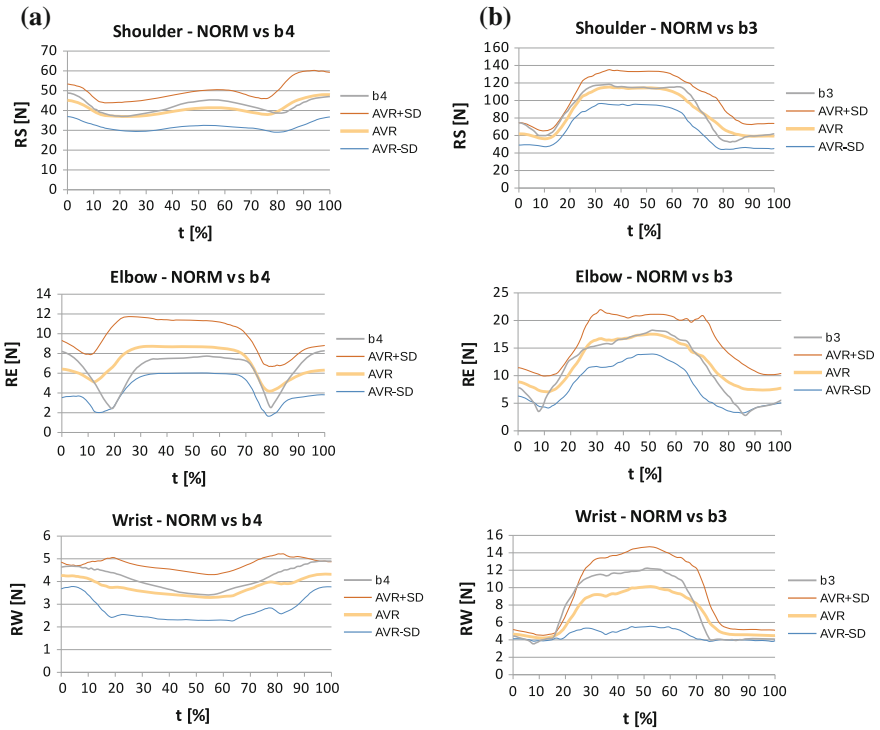


Fig. 4 Patterns of reaction forces in joints: **a** shoulder, elbow and wrist during *drinking from a cup* activity and courses of reaction forces of one of subject, **b** shoulder, elbow and wrist during *lifting of an object* activity and courses of reaction forces of one of subject

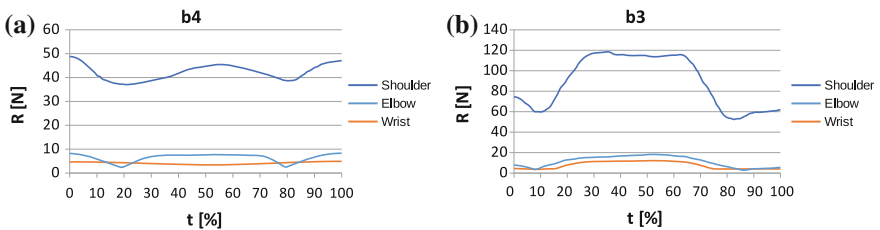
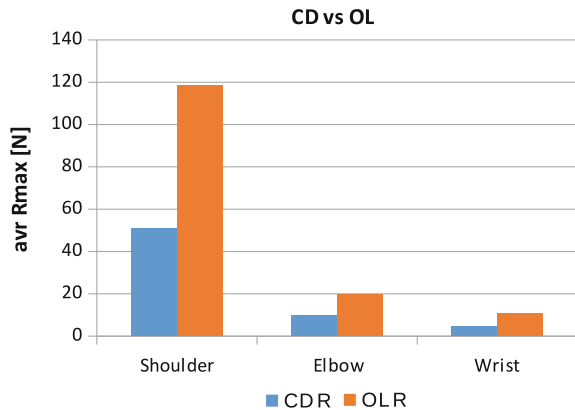


Fig. 5 Courses of force reactions in joints: **a** shoulder, elbow and wrist during *drinking from a cup* activity determined for one of examined man (b4), **b** shoulder, elbow and wrist during *lifting of an object* activity determined for one of examined man (b4)

was paid to the fact that not only muscles affect the results of calculations, but also the adjacent joints. The aim of their work was to create a three-dimensional model of the upper limb, which will include the major muscles of the upper limb. According to the analyses and the results it can be noticed that the average maximal reaction force in the shoulder joint during the procedure *lifting of an object* activity (OL)

Fig. 6 Comparison of average maximal forces reactions in joints: shoulder, elbow and wrist during *drinking from a cup* (CD) and *lifting of an object* (OL) activities



is about 131 % higher than the reaction occurring during *drinking from a cup* CD. A similar relationship was observed in the wrist joint where the average maximal reaction force in the joint during the *lifting of an object* activity (OL) is about 139 % higher than the force reaction occurring during *drinking from a cup* CD. The average maximal reaction force in the elbow joint during *lifting of an object* (OL) is about 104 % higher than the force reaction occurring during *drinking from a cup* (CD). The developed methodology of experimental and modelling research allows, among others, the determination of a number of parameters and courses of kinematic and dynamic parameters that describe the selected activities *drinking from a cup* and *lifting of an object* performed by the subject. The methodology applied in the research as in [8, 11, 17, 18] integrates experimental measurements of kinematic parameters, which are then used in numerical computation. A mathematical model of movement of the upper limbs allowing the determination of force reaction in joints in a non-invasive way while performing selected routine activities is an important element of the applied research methodology. The parameters achieved by this method, supplement the results obtained from experimental measurements. In the study the resultant reaction forces in the shoulder, elbow and wrist joints were determined and analyzed for the group of healthy people as well as patterns of joint reaction force of the individual joints were developed. The calculations of inverse dynamic problem (the carried out in the AnyBody programme were similar to Damsgaard [4] and they allowed the determination of reaction forces in the joints.

References

1. Albracht, K., Arampatzis, A., Baltzopoulos, V.: Assessment of muscle volume and physiological cross-sectional area of the human triceps surae muscle in vivo. *J. Biomech.* **41**, 2211–2218 (2008)
2. Andel C.J.A., Wolterbeek N., Doorenbosch C.A.M., Veeger D.J., Harlaar J.: Complete 3D kinematics of upper extremity functional tasks. *Gait Posture* **27**, 120–127 (2008) (Elsevier)

3. Chan A.M.A., Nicol A.C.: A new portable 3-D gyroscope system for the evaluation of upper limb function. In: ISB XXth Congress—ASB 29th Annual Meeting July 31–August 5, Cleveland, Ohio
4. Damsgaard, M., Rasmussen, J., Christensen, S.T., Surma, E., Zee, M.: Analysis of musculoskeletal systems in the anybody modeling system. *Simul. Model. Practice Theory* **14**, 1100–1111 (2006)
5. Gattamelata D., Pezzuti E., Valentini P.P.: Accurate geometrical constraints for the computer aided modelling of the human upper limb. *Comput. Aided Des.* **39**, 540–547 (2007) (Elsevier)
6. Guzik, A., Michnik, R., Rycerski, W.: The estimation of rehabilitation progress in patients with psychomotor diseases of upper limb based on modeling and experimental research. *Acta Bioeng. Biomech.* **8**(2), 79–87 (2006)
7. Guzik-Kopyto A., Michnik R., Wodarski P., Gzik M., Bieniek A.: Wyznaczenie obciążeń w układzie szkieletowo-mięśniowym kończyny górnej na podstawie badań kinematyki z wykorzystaniem inercyjnego systemu pomiarowego. (Determination of the load in the skeletal-muscle system on the basis of the kinematics using inertial measurement systems), *Modelowanie Inżynierskie* 2014 nr 53, tom 22, str. 62–69 (in Polish)
8. Holzbaaur, K., Murray, W.: A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control. *Ann. Biomed. Eng.* **33**(6), 829–840 (2005)
9. <http://homepage.tudelft.nl/g6u61/repository/shoulder/overview.htm>. Accessed 01 June 2013
10. http://wiki.anyscript.org/index.php/Main_Page. Accessed 22 Oct 2014
11. Ingram, M.A., Garth, J.R.: A study of external forces and moments at the shoulder and elbow while performing everyday tasks. *Clin. Biomech.* **19**, 586–594 (2004). Elsevier
12. Jurkojc, J., Michnik, R., Guzik-Kopyto, A., Gzik, M., Rycerski, M.: Kinematic differences in gait obtained for people with right and left paresis. *Information Technologies in Biomedicine*, pp. 464–471. Springer (2012)
13. Jurkojc, J., Michnik, R., Pauk, J.: Identification of muscle forces acting in lower limb with use planar and spatial mathematical model. *J. Vibroeng.* **11**(3), 566–570 (2009)
14. Klotz, M.C.M., Kost, L., Braatz, L.F., Ewerbeck, V., Heitzmann, D., Gantz, S., Dreher, T., Wolf, S.I.: Motion capture of the upper extremity during activities of daily living in patients with spastic hemiplegic cerebral palsy. *Gait Posture* **38**, 148–152 (2013)
15. Kim, K., Song, W.-K., Lee, J., Lee, H.-W., Park, D.S., Ko, B.-W., Kima, J.: Kinematic analysis of upper extremity movement during drinking in hemiplegic subjects. *Clin. Biomech.* **29**, 248–256 (2014) (Elsevier)
16. Magermans D.J., Chadwick E.K.J., Veeger H.E.J., Helm F.C.T.: Requirements for upper extremity motions during activities of daily living. *Clin. Biomech.* **20**, 591–599 (2005) (Elsevier)
17. Michnik, R., Jurkojc, J., Paszenda, Z., Michnik, A., Bachorz, M., Rycerski, W., Janota, J.: Application of motion analysis systems In the designing of rehabilitation device, *Information Technologies in Biomedicine*, vol. 2 Book Series: Advances in. Soft Computing **69**, 417–422 (2010)
18. Michnik, R., Jurkojc, J., Pauk, J.: Identification of muscles forces during gait of children with foot disabilities. *MECHANIKA* **6**, 48–51 (2009)
19. Michnik, R., Jurkojc, J., Rak, Z., Mężyk, A., Paszenda, Z., Rycerski, W., Janota, J., Brandt, J.: Kinematic Analysis of Complex Therapeutic Movements of the Upper Limbs. *Information Technologies in Biomedicine*. Book Series: Advances in. Soft Computing **47**, 551–558 (2008)
20. Murray I.A., Johnson G.R.: A study of the external forces and moments at the shoulder and elbow while performing every day tasks. *Clin. Biomech.* **19**, 586–594 (2004) (Elsevier)
21. Pennestri, E., Stefanelli, R., Valentini, P.P., Vita, L.: Virtual musculo-skeletal model for the biomechanical analysis of the upper limb. *J. Biomech.* **40**, 1350–1361 (2007)
22. Quental, C., Folgadoa, J., Ambrósioa, J., Monteiro, J.: Critical analysis of musculoskeletal modelling complexity in multibody biomechanical models of the upper limb. *Compu. Methods Biomech. Biomed. Eng.* **24** (2013)
23. Rab G., Petuskey K., Bagley A.: A method for determination of upper extremity kinematics. *Gait Posture* **15**, 113–119 (2002) (Elsevier)

24. Roman-Liu, D.: Tomasz Tokarski, T.: Upper limb strength in relation to upper limb posture. *Int. J. Ind. Ergon.* **35**, 19–31 (2005)
25. Vandenberghe, A., Levin, O., De Schutter, J., Swinnen, S., Jonkers, I.: Three-dimensional reaching tasks effect of reaching height and width on upper limb kinematics and muscle activity. *Gait Posture* **32**, 500–507 (2010). Elsevier
26. Westerhoff, P., Graichen, F., Bender, A., Halder, A., Beier, A., Rohlmann, A., Bergmann, G.: In vivo measurements of shoulder joint loads during activities of daily living. *J. Biomech.* **42**, 1840–1849 (2009)
27. Williams S., Schmidt R., Disselhorst-Klug C., Rau G.: An upper body model for the kinematical analysis of the joint chain of the human arm. *J. Biomech.* **39**, 2419–2429 (2006) (Elsevier)

Building Management System—Artificial Intelligence Elements in Ambient Living Driving and Ant Programming for Energy Saving—Alternative Approach

Jarosław Utracki

Abstract The paper describes problems of common fuzzy-logic based systems and offers a new approach. Genetic and ant-colony algorithms improvements to driving functions in Building Management Systems (BMS) are suggested. Two points of view at these systems are taken under consideration. First, from *outside* of specialization in designing energetically zero-emitting or low-energy buildings, this work postulates several changes in intelligent buildings design standards. Second, from *inside* look at Building Energy Management System (BEMS) energy consumption reduction, the main theme of this paper refers to the algorithms based on nature (ant programming and genetic algorithms) to solve big data processing problems. This paper is prepared to visualize the huge importance of preventing from non-intended Urban Heat Island Effect (UHIE) creation in human urbanized environment.

Keywords BMS · Intelligent buildings · UHIE · Genetic algorithms · Ant programming · Ant colony optimization · Travelling salesman problem

1 Introduction

Research purposes presented in the paper lead to examination of energy saving procedures formulated in alternative method as assisting supervisory system. The author's motivation is inspired by observation of the work of contemporary systems. It seems that present systems reach its maximum efficiency thus the author attempts to test known but not standard techniques in a new subject with irregular method. The author assumes that new procedures allow to provide subsequent order of magnitude energy saving. This approach is the path of study over Ant Colony Optimization (ACO) usability in real problems and different aspects of life. This thesis is yet to

J. Utracki (✉)

Institute of Computer Science, University of Silesia, Będzińska 39, 41-200
Sosnowiec, Poland
e-mail: Jaroslaw.Utracki@us.edu.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_10

be confirmed or rejected by experiments. The author has a strong conviction of the thesis' veracity due to the knowledge and executed observation after first simulations and preliminary tests.

2 Related Works

Literature studies have revealed numerous papers treating of ACO and the Travelling Salesman Problem (TSP) in the area of logistics. Multiple papers about Intelligent Energy Managing Systems using fuzzy-logic have been found, although the author seeks to narrow down the scope of search to articles referring to energy saving systems with ACO usage. No paper about the approach presented in this work has been found. Existence of such papers is unprovable, thus it can be assumed that corresponding materials are uncommon or represent an internal confidential work. The author does not have any knowledge of similar approaches. Works with ACO usage in the field of Intelligent Buildings yet in another meaning and implementation can be found in [1, 10, 17, 19, 28, 29].

3 Present State of BMS in Intelligence Buildings—Short Survey

Nowadays, buildings equipped with autonomic control systems are called *intelligent*. These systems may include the following components: power managing, interior climate managing (air conditioning, temperature), illumination managing, Closed Circuit TeleVision (CCTV), access and security control, etc.

The buildings' *intelligence* [25] may be various and generally depends on the architectural design, the developer, and the investor's plans and needs. Small buildings (e.g. places of residence) usually have one integrated circuit with Programmable Logic Controller (PLC) driver in closed technology for operation with single autonomous systems, such as air-conditioning, central heating, utility hot water, illumination, etc. (more information can be found in literature [21]). Such centralized systems may control behavior of window roller blinds, ventilators, car gates or grass irrigation and raise user's comfort without exact energy saving [3, 27].

Hybrid systems equipped with solar photovoltaics and solar heating systems stand as simple energy saving solution. These systems based on, e.g., KNX/EIB (world-wide standard for home and building control [4]) can be more efficient and flexible. The article constitutes an introduction to BMS in office buildings in which considerably larger order of magnitude systems can be managed, offering in total real power saving. This kind of buildings have enough expansive systems to justify implementation of decentralized BMS with multi-access and controlled by different aspects (access control, fire department, emergency, technical staff, etc.). Such systems are



Fig. 1 A building equipped with central and multi-role decomposition Energy Management System as an example of dispersed autonomous system—K1 (Błękitek, Kraków, Poland)

combined by several different single-function subsystems integrated together for improved (associated) management. The general concept is that every single room is controlled by sensors and its comfort parameters are driven by autonomic drivers (preparing usage circumstances by temperature, air-conditioning drivers), scheduling driving (changing eco-mode to human-comfort mode from non-usage period to normal human presence state by preordering in system scheduler) or manual driving (for non-automatic, occasional events) [26]. It appears to be managed from one central unit, as it was in such systems as HVAC (Heating, Ventilation, and Air Conditioning) [20], although it could be vulnerable in case of central unit failure. Due to the problem systems are designed as dispersed management systems (BDMS, Building Dispersed Management Systems) [13] such as in Fig. 1.

4 General Faults in Classic Concept of BEMS

Predominantly, Building Energy Management Systems can prove their efficiency in power saving. The issue does not represent a goal in this paper, therefore it can be assumed that such systems can reduce power utilization up to 80% [12, 15] due to existing implementation (measured as total energy utilization drawn from grid). The factor can only be increased through application of new building materials or renewable energy sources (solar panels, photovoltaics, wind generators, and so forth). Although these efforts do not reduce energy but merely refill power without external (public) electric grid usage and in effect, it reduces grid load.

Present executive drivers are based on self-learning algorithms with fuzzy-logic operability to provide efficacious work in varying environment with external logic circuit for scheduling (from central system), manual override or connection with extraneous weather prediction systems. The property of executive drivers is critical

for failure avoidance and determining reliability of the whole system. Central unit crash can prevent from scheduling or driving manually particular parts of managed building yet the whole system can work in autonomous mode with preprogrammed parameters and all subsystems can assure basic usability.

Conversely, what constitutes a benefit of this kind of systems, in the parallel manner is its fault. Particularly, cells of managed building (correlated with sensors in rooms) work to provide expected parameters in controlled area, although generally do not work in correlation together with surrounding drivers. It can lead to not optimal energy utilization in the whole system. Central unit do not has this role implemented either, limiting its function to supervision and delegation of expected parameters.

It can be noticed that this field requires improvement. For further process our domain of discourse will be narrowed to one subsystem for simplification of the goal of the paper. It will therefore focus on heating/cooling systems.

5 Ants Can Help

Ant Colony Programming (ACP) is an idea of using automatic programming for finding solutions to difficult or complicated problems for usual numeric evaluation. The broadly used programming techniques are Genetic Programming (GP) or Generic Algorithms (GA), which imitate phenomena existing in biology. The other technique is based on observation of ants' colony behavior, so it is also based on nature. More information about ACP can be obtained in literature [2, 5–7, 9, 11, 14, 16, 18, 22, 23].

5.1 *Idea of Postulate*

A typical heating system of a building contains pipes connected as a grid. The primary role of the grid is to transport heating or cooling factor to each area of building, where local drivers control internal temperature parameters. The heat is drawn from a central source (the cold is the result of lack of heating with natural energy dispersion to the colder environment). The system is controlled by a central managing unit which has two roles: statistical data collector and heat source moderator. Some rooms in this model will be overheated (internal rooms with small heat dispersion) or underheated (heating factor does not have enough energy after heat dispersion in previous rooms). Overheating can be controlled by an air-conditioning system. However, this leads to increase of the energy usage. Primarily, an extra dose of energy has been used to overheat and after that the energy has been used to cool down the rooms. As it can be noticed, the energy is wasted twice. The overdosed energy is commonly removed from building by spread to atmosphere. Air recuperation is one of the ways to minimize such a waste, yet it is not effective enough to avert energy expenditures. A proposed solution is to reorganize the way of heat delivery for all rooms in the

managed building in such a manner, that each room in building receives from a heating system precise dose of energy (no more than the room requires), which is demanded to provide desired exploitation parameters. However, it is impossible to achieve it in conventional heat grids.

5.2 Physical Model

Let us attempt to offer a physical model of higher quality than the existing ones, which can be applied in central supervising role. Let it be a matrix created in 3D similar to Fig. 2. The matrix has connections (vectors) between nodes in the grid.

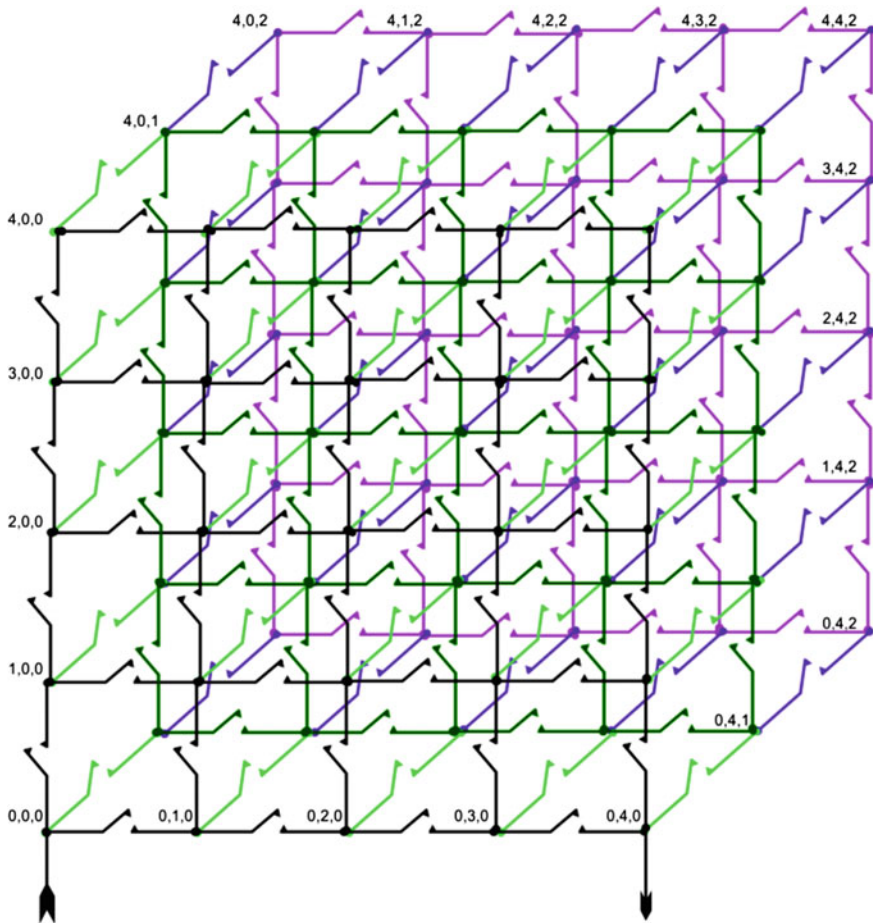


Fig. 2 Visionary (ideal) matrix

Each connection can be entirely transportable or blocked and each one can transport only in one direction at a time (directions can be switched). Transportation capacity is determined as a parameter of the vector between two nodes. Nodes have unique addresses (e.g. X/Y/Z position).

A connection can be traversed from the input point (XYZ:000) to every node in the matrix. Let us assume that there are workers W , where

$$W \in (1 \dots n) \tag{1}$$

and every worker can take package P consisting bubbles B , where

$$P \in (1 \dots m) \text{ and } B \in (1 \dots q) . \tag{2}$$

Each bubble vaporizes in vanishing time T_V , where T_V is greater than time of reach for the most distant node T_R . Each node has expectations for ordered quantity of bubbles present in time.

General activity is to deliver ordered quantity of bubbles to each node choosing the shortest way with a minimal bubble loss (vaporization). It is searched for a function describing expected activity. A possible real model is presented in Fig. 3.

It should be noticed that not every node is expected in a real building, as well quantity and position are reduced to real rooms—for approximation 1 node = 1 room. It is assumed that this is matrix with empty nodes.

The whole grid can be described node by node in a multi-pointer reference table with node notation presented in Fig. 4, where:

- NODE XYZ—ID/Address of the node,
- P₁ Addr—pointer to neighbouring node,

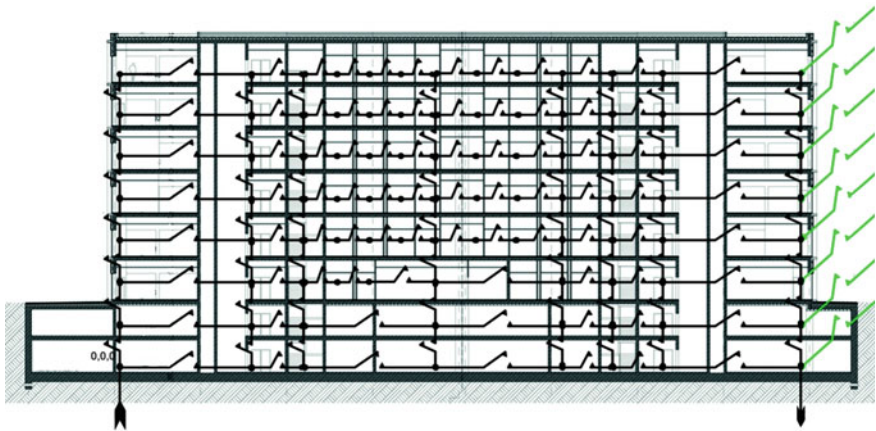


Fig. 3 Physical approximation of the matrix in a real building (in 2D for a preferable visibility—existing *green oblique lines* on the right remind that entire model is in 3D)

NODE XYZ	P₁	P₁	P₁	P₂	P₂	P₂	P_x	P_x	P_x
	Addr	Sem	Par	Addr	Sem	Par				Addr	Sem	Par

Fig. 4 Node’s references row

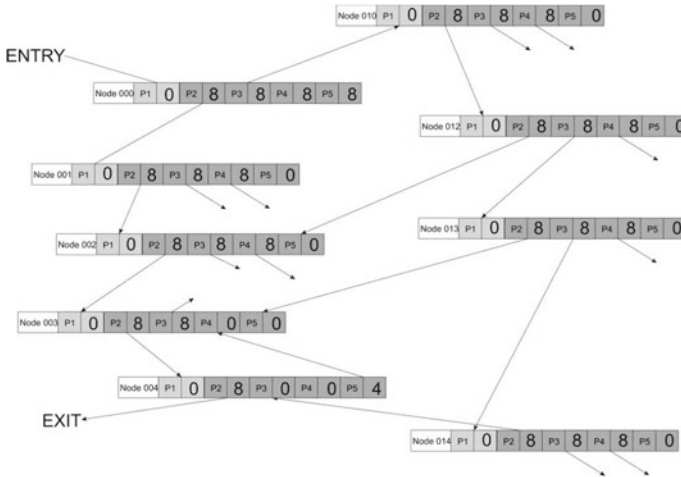


Fig. 5 Multi-pointer style vector map for physical grid

- P₁ Sem—semaphore for the connection with another node, where 0 means “no route” to node (blocked or reverse direction), value greater than 0 means transport capability of the connection,
- P₁ Par—additional parameters of particular connections, such as distance, bubble vaporization cost, etc.,
- P₂ . . . P_x—other P₂, P₃, . . . , P_x triples (P Addr : P Sem : P Par), where $x \in (2 \dots \infty)$, have similar meaning.

The whole physical grid can be described in this manner and a vector map can be received (Fig. 5).

In the real model the function describing expected activity is still being searched. It can be assumed that this function is highly complex and it cannot be implemented in Central Unit of BEMS in that shape. It is needed to find and approximate this function to the acceptable level of complexity.

5.3 Travelling Salesman Problem with ACO

The Travelling Salesman Problem is one of the primarily known NP-hard problems. It stands that there exist no exact algorithm solving it in polynomial time. The minimal expected time to obtain an optimal solution is exponential. Therefore, in many

implementations, an exact solution is not required and the approximated one can be obtained with heuristics and applied. Various algorithms can solve TSP more or less successfully [8].

5.4 *Modified TSP as the Solution of a Real Problem*

To solve our problem, the TSP modified algorithm can be indispensable. It constitutes the base for prepared experiments with simulated grid with workers-like model. As it can be noted, the workers are ants carrying bubbles (heat energy) and searching for the shortest (energetically) path to the nodes. When the Best Optimal Path (BOP) is found, a central managing system reorganizes its transport parameters, checks sensors (combined with PLC drivers) for updates and starts a new iteration.

5.5 *Experiment*

It can be assumed that heating grid controlled by ACO will be more efficient than the standard one based on fuzzy-logic drivers. An experiment has been conducted to verify this assumption. Three rooms in the simulated building equipped with central heating system and checking behavior of heating grid in 24 h time period have been chosen.

Basics:

- rooms: a lecture hall (room 'A'), a lobby (room 'B') and an office (room 'C'),
- programmed expected temperature for all the rooms,
- simulated external temperature.

After the experiment, impressive results have been received. The results for room 'A' are presented as temperature change plot for 24 h time period (Fig. 6).

The simulation has been conducted in three different modes:

- standard non-driven central heating grid (typical heaters equipped with thermostats),
- central heating with PLC drivers and onefold-way heat factor transport,
- central heating with PLC drivers, multi-way (three) ACO driven heat factor transport.

The preprogrammed and expected temperatures are presented by gray line P and external temperatures are represented by dotted black line Ex .

As it can be predicted, the first simulation (dotted line C_{ni}) is the worst. The temperature could not have been moderated during the whole period of time and a large dose of energy has been wasted on non-intended and unnecessary heating. This mode test has only an informational role and it can be omitted in a further evaluation.

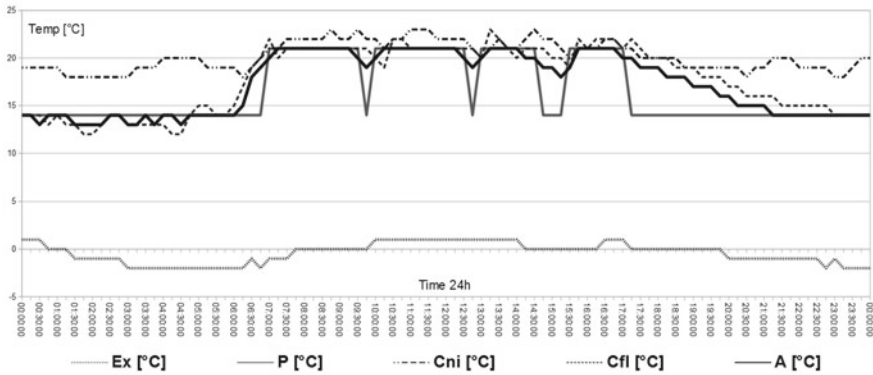


Fig. 6 Heating room ‘A’ during 24h time period

The second mode (dotted line C_{fl}): common, contemporary system with PLC shows that significantly large amount of energy can be saved.

The third one (thick black line A) is the result of the ACO and application of a multi-way heat factor transport in heating grid. As it can be noticed, the line is generally proximal to the preprogrammed temperature compared to the C_{fl} one. A short analysis reveals that C_{fl} is similar or almost the same to the preprogrammed temperature and delivers proper comfort of usage and a sufficient temperature. Still, small fluctuations and overheating can be recognized. Rooms ‘B’ and ‘C’ demonstrate similar behavior and have been omitted in further discussion.

All energy needed to ensure human comfort can be computed as the integral of the function:

$$C_{fL} = \int f(x) dx \cdot k. \tag{3}$$

And consequently, all energy required after ACO can be computed as:

$$A = \int p(x) dx \cdot k, \tag{4}$$

where k is constant factor. Both functions can be approximated from acquired data by ant-based function approximation or genetic algorithms function approximation either. If both functions are subtracted, the saved or wasted energy amount will be received, depending on sign of the result.

$$\Delta E = \left(\int f(x) dx - \int p(x) dx \right) \cdot k. \tag{5}$$

5.6 *Experiment Result Periphrasis*

The experiment data shows that $\approx 2.9\%$ of energy can be saved due to ant-colony optimization and multi-way transport grid. Using different grid or grid matrix reveals different results, so it would be possible to save larger or smaller amount of energy. The results can be determined through a multi-way feature, resembling a matrix-like grid. It can be a direction for further short and long-term examination with flexible models. Eventually, further work should lead to preparation of designed ant-colony based algorithms, grid's nodes notation for describing real grids (as an equivalent of the conventional notation for ant-colony based evaluation), software simulator testing usage of connections between nodes and reducing their number to minimal quantity by exclusion of unusable or rarely used ones, what would simplify complexity, lower costs and failure frequency of a grid.

6 **Problem with *Energy Refilling* by Renewable Energy Sources Instead of Power Reducing**

One of the disturbing thoughts is surplus heat spreading to the environment while exploiting a common building [24]. The renewable power sources can be applied and the zero-energy (passive) building can be obtained. The human comfort of living and use of maintenance in building leads to heat dump necessity when it is overheated. In the shortest manner it generates UHIE and leads to deterioration of living conditions outside a building. The postulated solution is application of ground buffer with heat pump to accumulate surplus heat in the basements of the building, use of rain water and reverse mode in cold seasons to heat a building with inconsiderable power amount instead of heat spreading.

7 **Conclusions**

It can be remarked that modified grid with fuzzy-logic PLC reached its order of magnitude savings and there do not exist "good" known methods to improve this problem. Further energy savings can be provided through suggested modification of system by additional central driving based on simulated grid with matrix connection in real grid. As the measurements revealed, about 2.9% of energy has been saved, although it may not be taken under consideration in small heating systems because of expensive and complex modification of a heating grid. Nevertheless, the conclusion of the experiment can be very promising for larger systems. Further experiments with simulated central heating systems standing as a base for ACP/ACO algorithms application are planned. Results of these experiments will be published in subsequent articles. General goals of the effort are to implement a full-size, real-working central heating system based on ACO in Intelligent Building for estimation of the natural conditions.

References

1. Amjad, A.-M., Hassan, M., Ashkan, R.K.: Optimal smart home energy management considering energy saving and a comfortable lifestyle. *IEEE Trans. Smart Grid* **6**(10), 324–332 (2015)
2. Bauer, A., Bullnheimer, B., Hartl, R.F., Strauss, Ch.: An ant colony optimization approach for the single machine total tardiness problem. Department of Management Science, University of Vienna, Austria. Proceedings: Congress on Evolutionary Computation (1999)
3. Borkowski, P.: *Inteligentne Systemy Zarządzania Budynkiem (Intelligent building management systems)*. Wydawnictwo Politechniki Łódzkiej 58–72 (2011) (in Polish)
4. Borkowski, P.: *Inteligentne Systemy Zarządzania Budynkiem (Intelligent building management systems)*. Wydawnictwo Politechniki Łódzkiej 7–47 (2011) (in Polish)
5. Boryczka, M., Boryczka, U.: Generative policies in ant system. Proceedings: Conference EUFIT'97, Aachen (1997)
6. Boryczka, M.: Ant colony programming for approximation problems. Proceedings: 11th International Symposium on Intelligent Information System, Sopot (2002)
7. Boryczka, M.: Programowanie Mrowiskowe w Procesie Aproksymacji Funkcji (Ant Programming in the Process of Approximation of Functions). Wydawnictwo Uniwersytetu Śląskiego 63–64 (2006) (in Polish)
8. Brezina, I., Čičková, Z.: Solving the travelling salesman problem using the ant colony optimization. *Manag. Inf. Syst.* **6**(4), 010–014 (2011)
9. Bullnheimer, B., Hartl, R.F., Strauss, Ch.: Applying the ant system to the vehicle routing problem. In: *Metaheuristic: Advances and Trends in local Search Paradigms for Optimization*, pp. 109–120. Kluwer Academic Publisher (1999)
10. Chandra Mohan, B., Baskaran, R.: A survey: ant colony optimization based recent research and implementation on several engineering domain. Department of Computer Science and Engineering, Anna University, Chennai, India. <http://www.sciencedirect.com/science/article/pii/S0957417411013996>. Accessed 15 Dec 2015
11. Doerner, K., Gronalt, M., Hartl, R.F., Reinmann, M., Strauss, Ch., Stummer, M.: Savings Ants for the Vehicle Routing Problem. Institute of Management Science, University of Vienna, Austria (2002)
12. Górka, A.: Instalacje Grzewcze w Budynkach Pasywnych – Podstawowe Różnice w Projektowaniu Architektoniczno-Budowlanym, Budynki Pełne Energii (Heating systems in passive buildings: Fundamental differences in architectural design and construction. Buildings full of energy). *Kalisz* 19–25 (2013) (in Polish)
13. Grad, P.: An idea of dispersed buildings management system. *Inteligentne budynki: Teoria i praktyka (Intelligent buildings: Theory and practice)* red. Mikulik, J, Oficyna wydawnicza Text, Kraków, pp. 127–146 (2010)
14. Green, J., Whalley, J.L., Johnson, C.G.: Automatic Programming with Ant Colony Optimization. University of Kent, Proceedings of the 2004 (2004)
15. Heim, D.: Standard Energooszczędny a Standard Pasywny – Podstawowe Różnice w Projektowaniu Architektoniczno-Budowlanym, Budynki Pełne Energii (Energy efficient standard and passive standard: Fundamental differences in the architectural design and construction. Buildings full of energy). *Kalisz* 19–25 (2013) (in Polish)
16. Hölldobler, B., Wilson, E.O.: *The Ants*. Springer, Berlin (1990)
17. Jinsung, Byun, Sehyun, Park: Development of a self-adapting intelligent system for building energy saving and context-aware smart services. *IEEE Trans. Consum. Electr.* **57**(1), 90–98 (2011)
18. Koza, J.R.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, a Bradford Book, pp. 6–8. The MIT Press, Cambridge (1992)
19. Lu, P., Chen, S., Zheng, Y.: Artificial intelligence in civil engineering, mathematical problems in engineering, vol. 2012. <http://dx.doi.org/10.1155/2012/145974>. Accessed 15 Dec 2015. Article ID 145974

20. Niezabitowska, E., Masły, D.: Office building quality according to development of intelligence building concept. *Inteligentne budynki: Teoria i praktyka (Intelligent buildings: Theory and practice)*, red. Mikulik, J., Oficyna wydawnicza Text, Kraków, p. 41 (2010)
21. Pacholski, K.: Systemy Zarządzania Inteligentnym Budynkiem, *Budynki Pełne Energii (Intelligent Building Management Systems, Buildings full of Energy)* Kalisz 19–25 (2013) (in Polish)
22. Roux, O., Fonlupt, C.: Ant programming: or how to use ants for automatic programming. In: ANTS'2000, pp. 121–129 (2000)
23. Sałustowicz, R.P., Schmidhuber, J.: probabilistic incremental program evolution: stochastic search through program space. In: *Machine Learning: ECML'97, Heidelberg, Berlin, vol. 1224*, pp. 213–220 (1997)
24. Solecki, W.D., Rosenzweig, C., Parshall, L., Pope, G., Clark, M., Cox, J., Wiencke, M.: Mitigation of the heat island effect in Urban New Jersey. *Glob.l Environ. Change Part B: Environ. Hazards* 6(1), 39–49 (2005)
25. Syposz, J., Jadwiszczak, P.: Zintegrowane Systemy Zarządzania Energią w Budynkach Biurowych (Integrated Energy Management Systems in Office Buildings). PAN Komitet Inżynierii Środowiska, Monografia nr 41, 9, 11–13 (2007) (in Polish)
26. Syposz, J., Jadwiszczak, P.: Zintegrowane Systemy Zarządzania Energią w Budynkach Biurowych (Integrated Energy Management Systems in Office Buildings). PAN Komitet Inżynierii Środowiska, Monografia nr 41, 9, 19–49 (2007) (in Polish)
27. Teletask System. <http://www.teletask.be>. Accessed 15 Dec 2015
28. Zhang, J., Tang, J., Zong, Q., Li, J.: Energy-saving scheduling strategy for elevator group control system based on ant colony optimization. In: *IEEE Youth Conference on Information Computing and Telecommunications (YC-ICT)*, pp. 37–40 (2010)
29. Zheng, Y., Chen, S., Lin, Y., Wang, W.-L.: Bio-inspired optimization of sustainable energy systems: a review. *Math. Probl. Eng.* 2013 (2013). <http://dx.doi.org/10.1155/2013/354523>. Accessed 15 Dec 2015. Article ID 354523

Part II

Bioinformatics

Application of Adjoint Sensitivity Analysis to Parameter Estimation of Age-Structured Model of Cell Cycle

Michał Jakubczak and Krzysztof Fajarewicz

Abstract Age-structured models describe heterogeneous populations of individuals. Life of the individual consist of a finite number of phases, which differ in properties such as the ability to reproduce. This approach allows more realistic modelling of the population growth than using models assuming homogeneity of the population, which is especially important for complex organisms with long reproduction time. The aim of this study was to develop a methodology to estimate parameters of the age-structured model of cell cycle using adjoint sensitivity analysis. Results were obtained for artificially generated input data.

Keywords Age-structured models · Adjoint sensitivity analysis · Parameter estimation · Cell cycle

1 Introduction

In the classical population models, like Malthusian growth model or Lotka-Volterra equations [1], populations are homogeneous, i.e. all individuals are the same. This approach assumes that newly born individual is capable of reproduction, what is not true for species with relatively long life time, like humans. For better description of such species it is possible to introduce phases of life of individuals: an early phase of adolescence, a phase with ability to reproduce and a phase of senility. What is important, phases continues in a specific order, creating an age structure in the population.

In this work we introduce a novel method for parameter estimation in age-structured models. Our method is based upon a structural adjoint sensitivity analysis,

M. Jakubczak (✉) · K. Fajarewicz
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Automatic Control, Silesian University of Technology,
Akademicka 16, Gliwice, Poland
e-mail: Michal.Jakubczak@polsl.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_11

which is especially useful for models described by the block diagrams and was originally developed for neural networks [4] and afterwards was used for models of cell signalling pathways [5, 6] and systems with delays [7].

The main goal of the work was to create an adjoint sensitivity analysis methodology. Since the age-structured model parameter estimation is complex enough, we decide to use a simple cell cycle model, which allows us to avoid the model influence on the obtained results. The model can be easily modified to the age-structured form. It is also easy to gather experimental data for this model, because cells division takes approximately 24 h [10].

2 Model of a Cell Cycle

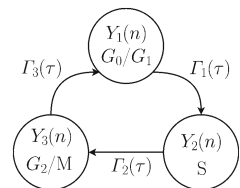
The model of cell cycle used in this work was proposed in [8]. It is composed of three phases: combined G_0/G_1 , S and combined G_2/M . Newly divided cells go to the G_0/G_1 phase, where they can start a new cycle by entering G_1 growth phase or stay in G_0 phase and wait for better external conditions. The software developed in this work is designed to work on the data from the flow cytometry and therefore it is impossible to separate these two phases. During the next phase of the model, the S phase, DNA in cells is replicated. After that, cells enter the G_2 gap phase and prepare for mitosis, which occurs in M phase. These two phases are also indistinguishable using the flow cytometry. Mitosis is the final phase of the cell cycle and it ends with the cytokinesis, which produces two cells from the one.

There is an age-structure of a subpopulation $Y_i(n)$ in each phase, i.e. individuals in the subpopulation differ in the dwell time τ in the phase. In every moment of global moment of time n the individual can go to the beginning of a next phase, become older in a current phase or die—these events depends on a transition functions $\Gamma_i(\tau)$, which determine their probability. Figure 1 presents the basic structure of the model.

Mathematical Description: The model is described by 3 difference equations:

$$\begin{aligned} Y_1(n+1) &= 2B_3Y_3(n) + A_1Y_1(n), \\ Y_2(n+1) &= B_1Y_1(n) + A_2Y_2(n), \\ Y_3(n+1) &= B_2Y_2(n) + A_3Y_3(n), \end{aligned} \quad (1)$$

Fig. 1 A structure of the cell cycle model used in the work



where:

$$Y_i(n) = [y_1^i(n), y_2^i(n), \dots, y_k^i(n)]^T, \tag{2}$$

$$A_i = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ s_1^i & 0 & \dots & \dots & \vdots \\ 0 & s_2^i & \ddots & \dots & \vdots \\ \dots & 0 & \ddots & 0 & \vdots \\ 0 & \dots & 0 & s_{k-1}^i & 0 \end{bmatrix} \quad B_i = \begin{bmatrix} r_1^i & r_2^i & \dots & \dots & r_k^i \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix} \tag{3}$$

$Y_i(n)$ is the vector of subpopulation of i -th phase and y_τ^i is the number of cells in i -th phase with age τ . k is the maximum τ per single phase. Matrices A_i are responsible for cells aging and matrices B_i correspond to cells transition to the next phase. Cell death is omitted. Both matrix types contain information about transition functions:

$$s_\tau^i = 1 - \Gamma_i(\tau), \tag{4}$$

$$r_\tau^i = \Gamma_i(\tau). \tag{5}$$

The model can be written as a block diagram (Fig. 2), what will be useful in the further analysis.

Parameters of the Model: Parameters in this case are transition functions (TFs) and initial cell distributions (ICDs) $f_i(\tau)$ in each phase. TFs are constant in every moment of time [3] and their value depends on τ . In our work we assumed TFs to be cumulative distribution functions, i.e. longer dwell time in the phase results in higher probability to go to the beginning of the next phase.

Total initial number of cells in each phase is known, so the second class of parameters are only ICDs given as probability density functions.

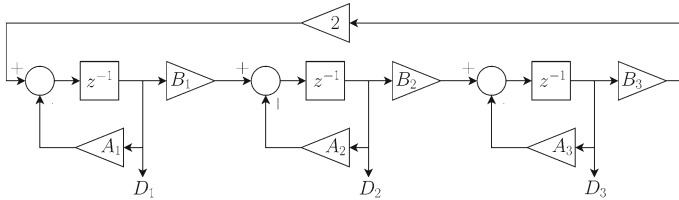


Fig. 2 Structural diagram of the cell cycle model. $D_i(n)$ are vectors of cell number in the i -th phase in the n -th moment of time, used as a reference data in parameter estimation with artificially generated data

3 Adjoint Sensitivity Analysis

Parameter estimation is performed by a gradient minimisation of index given as:

$$J = \frac{1}{2} \sum_{n=1}^N \sum_{i=1}^K (Y_i^{total}(n) - D_i^{total}(n))^2 \rightarrow \min, \tag{6}$$

where $D_i(n)$ is reference data and $Y_i(n)$ is the output from the model.

For an effective search through a parameter space for an optimal set of parameters it is necessary to calculate gradients of performance index in respect to these parameters. This task is not trivial for most complicated models, however, a model given by structural diagram can be easily modified to a form, which outputs the desired gradients [4, 6]. First of all, an expanded model based on the original model must be build (Fig. 3).

The expanded model is the base for building sensitivity and adjoint models. The aim is to calculate all sensitivity coefficients (sensitivities) given as:

$$s_i(u_{nom}) = \left. \frac{\partial J}{\partial u_i} \right|_{nom}, \tag{7}$$

where u_i is the i -th input and u_{nom} is a nominal input vector.

There are two ways to calculate all sensitivities in this case: using forward sensitivity analysis or adjoint sensitivity analysis. If number of inputs is greater than number of outputs it is better to use the adjoint approach due to lower calculation complexity. In the expanded model on Fig. 3 there is only 1 scalar output (performance index)

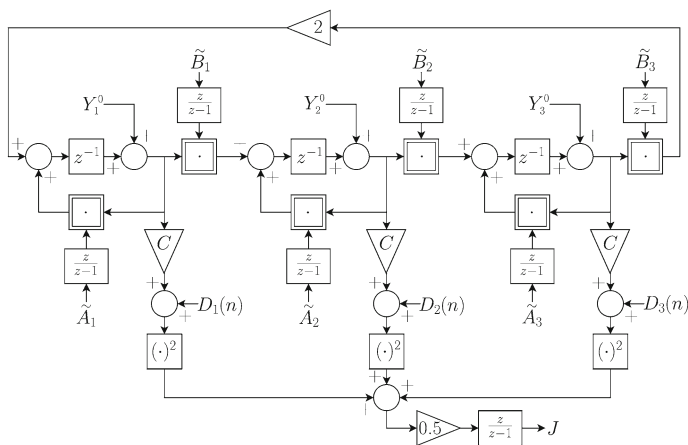


Fig. 3 Expanded model with matrices A_i and B_i and initial cell distributions vectors Y_i^0 as inputs and the performance index J as an output. ICD vectors are non-zero only in the initial moment of time, i.e. for $n = 0$

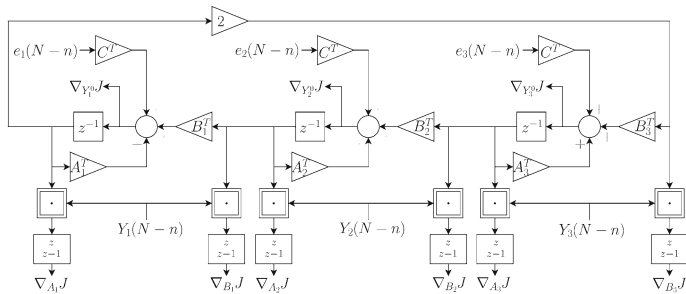


Fig. 4 Constructed adjoint system used to calculate all the sensitivity coefficients. $e_i(N - n) = Y_i^{total}(N - n) - D_i^{total}(N - n)$

and 9 multidimensional inputs. Because of that fact, we constructed the adjoint system using specific rules described in [4, 6]. The final system for the gradients calculation is presented in Fig. 4.

4 Parameter Estimation

Non-Parametric Approach In the non-parametric approach every single value of the TFs and ICDs were calculated separately. Because information about values of the TFs is included in matrices A_i and B_i (3), then final gradients of the performance index in respect to TFs are given by:

$$\nabla_{\Gamma_i(\tau)} J = \nabla_{r_i^1} J - \nabla_{s_i^1} J. \tag{8}$$

The gradients of the performance index in respect to the initial cell distributions are given directly from the adjoint system.

Parametric Approach The parametric approach assumes that parameters are described by some functions. In our work we proposed logistic growth [9] for the TFs (9) and normal probability density function for the ICDs (10).

$$\Gamma_i(\tau) = \frac{c_i}{1 + e^{-a_i\tau + b_i}} \tag{9}$$

$$f_i(\tau) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{(\tau-\mu_i)^2}{2\sigma_i^2}} \tag{10}$$

Gradients in respect to the parameters in Eqs. 9 and 10 are given by:

$$\nabla_{p_m} J = \sum_{j=1}^k (\nabla_{g_j^i(\tau)} J \cdot \nabla_{p_m} h_1), \tag{11}$$

where g is the TF or ICD and p_m is the m -th parameter of the assumed function h .

The advantage of the parametric approach is that much less parameters have to be estimated than in the non-parametric approach. A single TF in this case is described by only 3 parameters and ICD by 2 parameters. What is important, these parameter numbers are independent of maximum τ per single phase (k)—in the non-parametric approach there is k parameters which have to be estimated per one TF or ICD. In the other hand, incorrect assumption of the function's form may lead to incorrect results.

Calculated gradients were used to the minimisation of the performance index executed by Matlab[®] function *fmincon*. The range of values for the transition functions was set as $(0, 1)$.

5 Results and Discussion

Reference data for the parameter estimation were generated with the basic model (Fig. 2). Results of estimation are presented in Fig. 5 (non-parametric approach) and Fig. 6 (parametric approach).

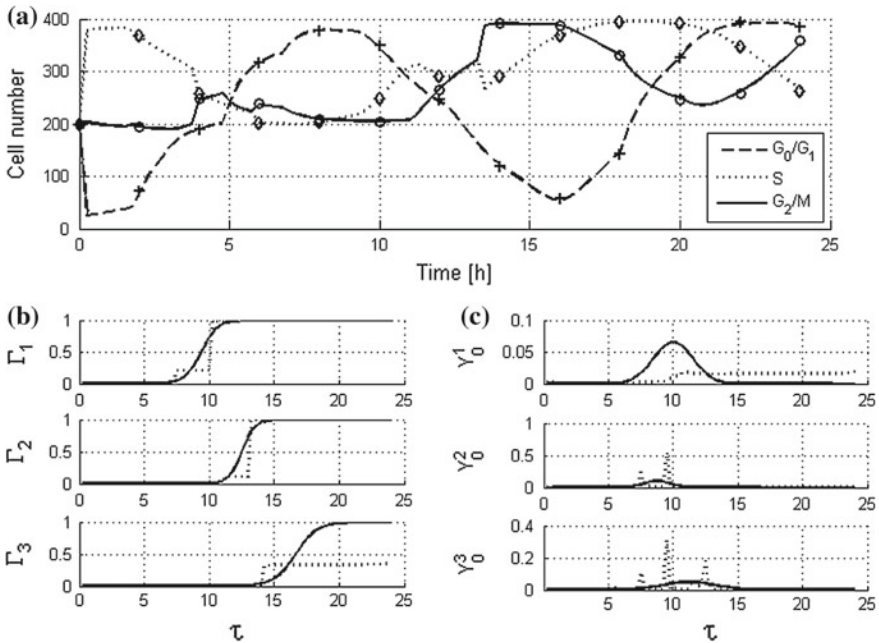


Fig. 5 Results for the parameter estimation using the non-parametric approach and artificial data. $J = 134.23$; $MSE_r = 4.45 \cdot 10^{-2}$; $MSE_f = 1.66 \cdot 10^{-3}$ **a** Fit of the model (lines) to the input data (+ for G_0/G_1 , \diamond for S , \circ for G_2/M). **b** TFs: estimated (dotted) and real (solid). **c** ICDs: estimated (dotted) and real (solid)

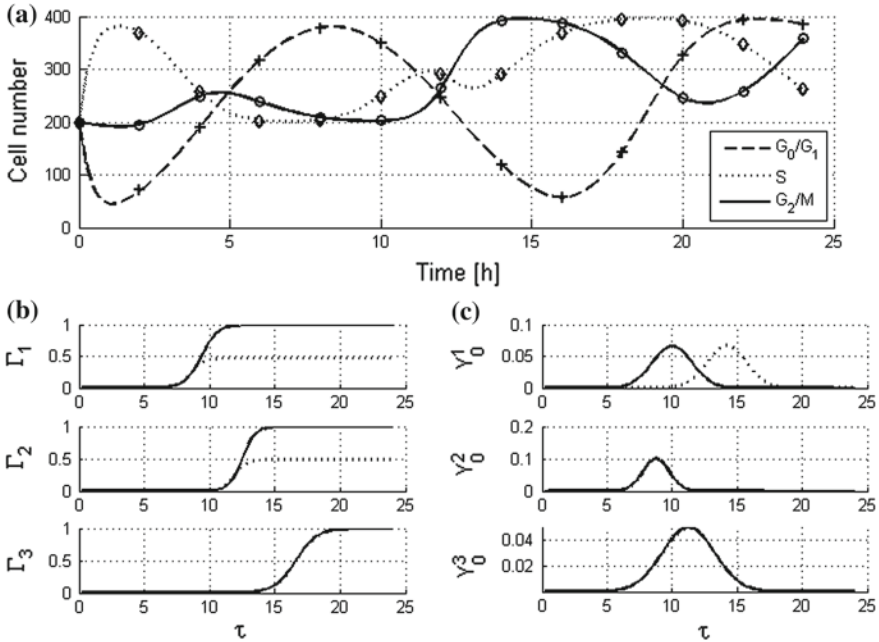


Fig. 6 Results for the parameter estimation using parametric approach and artificial data. $J = 9.27$; $MSE_r = 9.34 \cdot 10^{-2}$; $MSE_f = 2.81 \cdot 10^{-4}$ **a** Fit of the model (lines) to the input data (+ for G_0/G_1 , \diamond for S , \circ for G_2/M). **b** TFs: estimated (dotted) and real (solid). **c** ICDs: estimated (dotted) and real (solid)

Mean squared errors (MSE) were calculated to evaluate the fit of the estimated parameters to their real values:

$$MSE_g = \frac{1}{Kk} \sum_{i=1}^K \sum_{j=1}^k (\hat{g}_j^i - g_j^i)^2. \tag{12}$$

A very good fit of the cell number time courses to the input data with both approaches was obtained. Slightly better MSE results were achieved for the parametric approach, what is related to the proper parameter functions assumptions.

Estimated transition functions in both approaches introduce a cell death to the model ($\Gamma_{\tau_{max}}^i < 1$)—cells with the dwell time greater than τ_{max} are “falling out” of the phase without going to the beginning of the next one. It is probably related to the deterministic nature of the model. The assumed TFs (as the cumulative distribution functions) may not correspond to the real forms of distributions.

In the non-parametric approach TFs seems to have step forms. It is probably caused by the applied boundaries to the values of the TFs—the minimisation algorithm couldn’t find the mild form of the functions. This causes that the dwell time is almost the same for all cells in the particular phase.

The ICDs in the parametric approach were estimated almost perfectly—only the ICD in the G_0/G_1 phase is slightly shifted, however, a shape of the distribution is preserved. In the non-parametric approach, the estimated ICDs are irregular—it is caused by the same factor as in the TFs estimation.

6 Conclusions

A good fit of the model to the input data can be obtained using more than one set of parameters, i.e. estimated parameters providing a very good fit may not correspond to their real values. A gradient minimisation using adjoint sensitivity analysis is an effective tool to perform a parameter estimation of the age-structured models. The main advantage of the parameter estimation method used in our work is automation of the process. We are aware that this method requires longer computing time and is less accurate than other techniques based on regression [2] or Bayesian methods [11], however, it can be applied for any model introduced as a block diagram.

The comparison of the method used in our work with other parameter estimation methods will be under further investigations of our work. We will also examine another forms of the parameters used in the parametric approach and will use experimental input data to test our methodology in reality.

Acknowledgments The presented work was supported by the Polish National Science Center (NCN) under grants: DEC-2013/11/B/ST7/01713 (MJ) and DEC-2012/05/B/ST6/03472 (KF).

References

1. Albrecht, F., Gatzke, H., Haddad, A., Wax, N.: The dynamics of two interacting populations. *J. Math. Anal. Appl.* **46**, 658–670 (1974)
2. Deriso, R.B.: Harvesting strategies and parameter estimation for an age-structured model. *Can. J. Fish. Aquat. Sci.* **37**(2), 268–282 (1980)
3. Dyson, J., et al.: A nonlinear age and maturity structured model of population dynamics. *Basic theory. J. Math. Anal. Appl.* **242**, 93–104 (2000)
4. Fajarewicz, K., Galuszka, A.: Generalized Backpropagation Through Time for Continuous Time Neural Networks and Discrete Time Measurements. *Lecture Notes in Computer Science*, vol. 3070, pp. 190–196. Springer (2004)
5. Fajarewicz, K., Kimmel, M., Swierniak, A.: On fitting of mathematical models of cell signaling pathways using adjoint systems. *Math. Biosci. Eng.* **2**(3), 527–534 (2005)
6. Fajarewicz, K., Kimmel, M., Lipniacki, T., Swierniak, A.: Adjoint systems for models of cell signalling pathways and their application to parameter fitting. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **4**(3), 322–335 (2007)
7. Fajarewicz, K., Lakomic, K.: Parameter estimation of systems with delays via structural sensitivity analysis. *Discret. Contin. Dyn. Syst.-Ser. B* **19**(8), 2521–2533 (2014)
8. Sherer, E., et al.: Identification of age-structured models: cell cycle phase transitions. *Biotechnol. Bioeng.* **99**, 960–974 (2008)
9. Steel, G.G.: *Growth Kinetics of Tumours: Cell Population Kinetics in Relation to the Growth and Treatment Cancer*, p. 351. Clarendon Press, Oxford (1977)

10. Vermeulen, K.: The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif.* **36**, 131–149 (2003)
11. Walters, C., Ludwig, D.: Calculation of Bayes posterior probability distributions for key population parameters. *Can. J. Fish. Aquat. Sci.* **51**(3), 713–722 (1994)

Multidimensional Correlated Mutation Analysis for Protein Contact Map Prediction

Kristyna Kupkova, Karel Sedlar and Ivo Provaznik

Abstract Correlated mutation (CM) analysis has been proved to be an important tool used in protein contact map prediction from primary amino acid sequence. Over the last years CM methods have been refined and then often combined with methods of different nature in order to improve reached precision. However, since the methods are still relatively new, only precision and improvement have been reported without mentioning recall which is very low. In this paper, we combine previously described CM analysis methods with an outcome of four new techniques which significantly increase recall for the cost of minimal precision impairment. This study is the first of our knowledge with focus on recall improvement.

Keywords Correlated mutations · Contact map · Protein structure prediction · Residue contact

1 Introduction

The structure of a protein is one of the most important information carriers in molecular biology. Currently, there are several methods used to identify the three-dimensional (3D) architecture of a protein, including X-ray crystallography, NMR spectroscopy, and electron microscopy. These are, however, still expensive and time-consuming. Therefore, several computational methods have been recently developed, to predict residue contacts straight from primary amino acid sequence [5–7, 10, 16]. Correlated mutations (CM) analysis is one of the most promising approaches using multiple sequence alignment (MSA) of homologous proteins to identify residues that are in close proximity within the 3D structure, but distant in the primary structure.

MSA consists of highly conserved positions along with positions which contain mutations of varying scope. Not only the conserved positions, but also the ones with

K. Kupkova (✉) · K. Sedlar · I. Provaznik
Department of Biomedical Engineering, Brno University of Technology,
Technická 12, CZ 616 00 Brno, Czech Republic
e-mail: xkupko00@stud.feec.vutbr.cz

various mutations carry information about the protein structure and function. How proteins preserve their structure and function despite spontaneous mutations acquired due to evolution is the foundation of CM analysis, as in order for protein to function properly, it is important that negative point mutations are being compensated for by other mutations [7]. Therefore, the amino acid positions of vital importance to the protein folding have to evolve in synchrony which consequently results in a high correlation score.

Several methods, which use CM to predict residue contacts, have been developed. These methods usually include also another approach apart from a CM analysis in order to increase their precision. Alexov and co-workers reported method of CM analysis without using correlation score, as usually, but instead defined four parameters which are subjects of optimization [10]. After applying a filter defined by physical-chemical properties to set of 65 proteins, the overall reached precision was 0.09, with improvement 30 %. Casadio and co-workers upgraded CM analysis by combining it with neural networks [5]. Using the method on set of 173 non-homologous proteins resulted in 0.21 precision with an improvement >6. Apart from structure prediction, CM analysis has been proven to be an important tool for analysis of protein function, and of protein interaction networks, e.g. prediction of γ -shaped and θ -shaped networks in catalytic domains of Ser/Thr protein kinases by Feng, Long and co-workers [19] or study of HIV-protease evolution by Hamacher [9].

In previous studies, precision and improvement were the main parameters used to evaluate the success of the relevant method without mentioning recall which decreases rapidly with increasing precision. Thus the main goal of our study is to combine several CM analysis methods in order to increase recall with minimal decrease of precision. We use in total three methods to compute CM: mutual information (MI) [3] along with residue correlation (RC) analysis [7], and Jaccard index (JI) [11]. These methods are then combined with each other resulting in increased recall to various extent. Furthermore we used previously introduced filters reflecting physical-chemical properties of amino acids [10] in faith of increasing previously decreased precision values without again decreasing recall. However in average this step only led to immense recall impairment without significant precision amelioration.

2 Materials and Methods

2.1 Sequence Collection and Pretreatment

Each set of homologous protein chains was obtained by subjecting the protein chain of interest to PSI-Blast (Position-Specific Iterated Blast) run against the database of non-redundant protein sequences obtained from the National Center for Biotechnology Information (NCBI). Both orthologs and paralogs were included in the set, since it was demonstrated that this approach can lead to an increase of precision in

Table 1 Dataset specification

PDB ID	Chain ID	Seq length (bp)	MSA width	MSA depth
1A6K	A	151	151	61
1AFC	A	140	106	24
1AX8	A	166	138	36
1DKG	A	197	120	27
1GD1	O	334	288	204
1G9L	A	144	136	61
1HM6	A	346	337	51
1XQ5	A	143	128	107
2RQB	B	83	77	69
3BCQ	A	143	130	88
4DT6	A	240	215	69
4FFJ	A	210	158	92
1C27	A	263	181	48
1EXR	A	148	118	73
1PKQ	E	138	113	22
5E29	A	142	126	92

correlated mutation based analysis [1]. Only the sequences sharing more than 70 % and less than 90 % of identity with the reference sequence were selected. In case, that this step resulted in set of less than 125 sequences [12], the threshold was lowered to 65 %. Resulting sequences were aligned using Clustal Omega program [17]. The formed MSA was further processed by removing sequences having more than 90 % identity with each other in order to reduce covariation caused by phylogenetic influence of closely related sequences [1]. It is crucial to keep the reference sequence in the MSA during MSA purgation. Further MSA treatment included removing totally conserved positions, for it does not provide information relevant to CM analysis [1]. Also positions with gap occurrence in the reference sequence were excluded, which allows us to compare the final outcome with result obtained from PDB file of the reference sequence. Ultimately the cases with final MSA depth lower than 20 sequences were removed from protocol as these are considered not to be carrying enough information [10]. The final 16 sets of proteins along with characteristics of their MSAs are listed in Table 1.

2.2 Mutual Information

The mutual information (MI), based on Shannon's entropy, is a measure of reduction of uncertainty. MI between two columns reflects, how correlated is the pattern between these two columns [12], and is defined as [2, 3]:

$$MI_{ij} = \sum_{x \in X_i} \sum_{y \in Y_j} p_{ij}(x, y) \cdot \log_{20} \left(\frac{p_{ij}(x, y)}{p_i(x)p_j(y)} \right) = H_i(x) + H_j(y) - H_{ij}(x, y), \quad (1)$$

where x and y are the outcomes for random variables X_i and Y_j , drawn from a standard amino acid symbol set containing symbols for all 20 amino acids, taken from columns i and j of MSA. Variables $p_i(x)$, $p_j(y)$ represent the observed frequencies of occurrence of amino acids x and y in columns i and j of MSA and $p_{ij}(x, y)$ stands for joint probability function. $H_i(x)$ and $H_j(y)$ are column entropies and $H_{ij}(x, y)$ is the joint entropy. Entropies are calculated using a \log_{20} scale so that their values lie within the range of 0–1.

In order to lower background MI caused by finite sample size effect, and by phylogenetic influence, MI needs to be normalized. We used $MI/H_{ij}(x, y)$ normalization which was proved to have the best performance [12].

2.3 Residue Correlation Analysis

Correlation coefficient between two residues, i.e. the inner product between two unit vectors of length N^2 is computed as [7]:

$$r_{ij} = \frac{1}{N^2} \sum_{kl} \frac{w_{kl}(s_{ikl} - \langle s_i \rangle)(s_{jkl} - \langle s_j \rangle)}{\sigma_i \sigma_j}, \quad (2)$$

where values s_{ikl} , s_{jkl} are taken from mutation matrices. The mutation matrix $s(i, k, l)$ is composed of residue-residue distances of all possible pairs of sequences at position i in MSA, where k and l are the indexes of sequences running from 1 to N . We used BLOSUM62 substitution matrix as a measure of residue-residue distances, for superiority of McLachlan matrix, which was used in previous studies [5], was not justified in literature and provides similar results to the ones obtained from other substitution matrices [14]. Gap similarity was set to dummy value of 0 [5]. The value of σ_i is the standard deviation of s_{ikl} about the mean s_i . Weights w_{kl} are used to down-weight the influence of very similar sequences. The weight value for pair of sequences k and l is defined as fraction of amino acid mismatches over the alignment length L [15]:

$$w_{kl} = 1 - \frac{1}{L} \sum_i^L \delta(R_{ik} R_{il}). \quad (3)$$

Weights are further normalized to sum to 1 [7]. The final values of correlation coefficient r were also normalized to lie within the range of 0–1 by using: $r_{ij} = (r_{ij} - \min(r)) / (\max(r) - \min(r))$.

2.4 Jaccard Index

Jaccard index which is another measure of CM, is defined as [11]:

$$JI_{ij} = \frac{N_{xy}}{N_{xy} + N_{x0} + N_{y0}}, \quad (4)$$

where N_{xy} is the number of sequences, in which X_i and Y_j are jointly mutated, N_{x0} represents the number of sequences, with only X_i mutation, but not Y_j mutation and N_{y0} is the number of sequences with only Y_j mutation, but not X_i mutation.

2.5 Multidimensional CM Methods

As said before, the contact positions are expected to be the ones with a high correlation score, however each method described above is different, and brings different results. Therefore we set an imaginary multidimensional space, where each dimension is represented by correlation score of given method, and where a combination of positions represented by e.g. low MI value can still be in contact if the RC or JI values are high. And since the main goal of this study is to increase the number of correctly detected contacts, and therefore to increase recall, the number of true predictions out of all existing contacts ($recall = TP / (TP + FN)$), this idea seems like the most convenient solution how to reach desired results.

In our approach, predictions are not obtained by simple disjunction of CM methods results, but cutoff values are used to define first quadrant of an imaginary ellipse in 2D space (if two CM methods are combined) or part of an imaginary ellipsoid in 3D space (if three CM methods are combined). Two positions are then considered to be in contact if the following criteria are met:

$$\frac{MI_{ij}^2}{MI_{cut}^2} + \frac{RC_{ij}^2}{RC_{cut}^2} + \frac{JI_{ij}^2}{JI_{cut}^2} > 1. \quad (5)$$

Here MI_{cut} , RC_{cut} , and JI_{cut} stand for cutoff values of MI, RC and JI. The criteria introduced in (5) represent the case, when all MI, RC, and JI represent the dimensions of one multidimensional CM method (MRJ). In case that selection of only two methods is combined into one method, the correlation scores of the third method are omitted from the criteria. In final, this results in four new CM analysis methods: MRJ (combination of MI, RC, and JI), MR (combination of MI, RC), MJ (combination of MI, JI), and RJ (combination of RC, JI).

2.6 Filtration

The results obtained from methods mentioned above are further filtered by use of rules defined in [10], where energetically favorable contacts are those who form either a hydrophobic pair, an ion pair, a disulfide bridge, a hydrogen bond or pairs, in which donation of a hydrogen bond to a hydrogen acceptor is possible. Gly-Gly pair does not fall into any of categories mentioned above, but since this amino acid combination was frequently observed in residue pair contacts, it was added to list of admitted pair predictions. On the contrary, Trp-Phe, Phe-Ile, Trp-Leu, Trp-Pro, Pro-Ile, Pro-Leu, Pro-Met and Pro-Trp pairs were observed only sparsely and hence excluded from the list.

2.7 Contact Map

Protein can be represented by $N \times N$ symmetrical matrix (for protein of N atoms), where values at corresponding indexes represent distances among atoms in the protein. Contact map is an alternative binary representation of a distance map. It is also a symmetrical $N \times N$ matrix (here N represents number of residues), however, elements of this matrix are represented by only values 1 or 0 depending on whether the pair of residues is or is not in contact [5]. Two residues i and j in contact matrix C are considered to be in contact if the distance between them $\delta(i, j)$ is lower than a certain threshold μ [8]:

$$C_{ij} = \begin{cases} 1, & \delta(i, j) < \mu \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

The threshold values vary among different authors, e.g. in [8] is the cutoff distance defined as 6–7 Å, in [18] is the value 8 Å. Also the distance computation definition can be different among different authors, in [4] the distance for contact determination is the distance between $C_\alpha - C_\alpha$ atoms, in [13] it is the distance between $C_\beta - C_\beta$ atoms or another interpretation is the minimal distance between atoms belonging to the side chain or to the backbone of the two residues [5]. In this study we use the definition of a contact defined in [13], where the distance between two residues is computed as the distance between $C_\beta - C_\beta$ atoms and the value of threshold is set to 8 Å.

3 Results

3.1 Cutoff Specification

Random sample of 8 proteins (half of the whole dataset) was selected for cutoff value specification of MI, RC, and JI. Every cutoff value within a range of 0–1

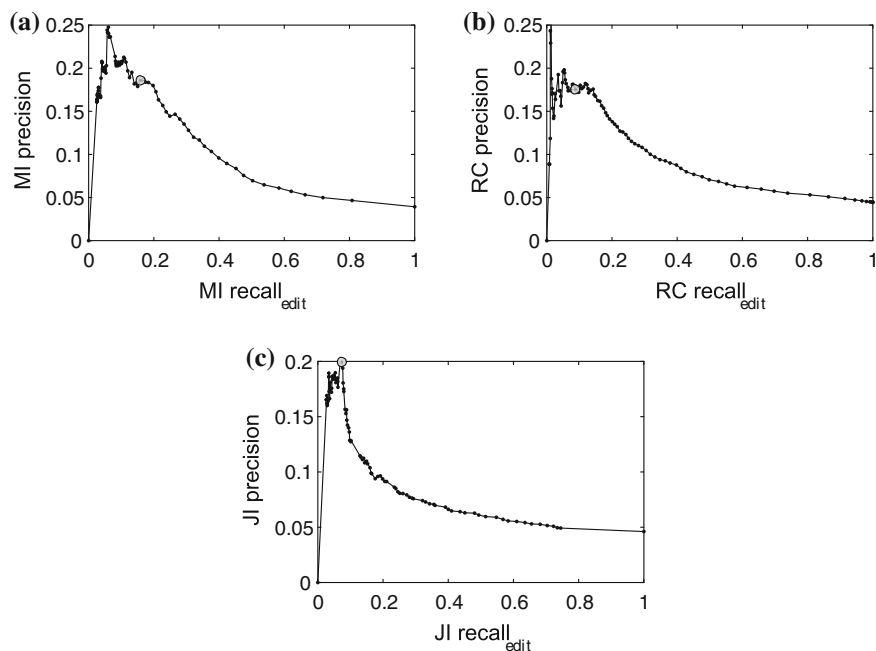


Fig. 1 Averaged precision-recall curves of 8 proteins. Each graph represents the characteristics of the contact prediction using values of **a** MI > MI cutoff, **b** RC > RC cutoff, **c** JI > JI cutoff. Precision-recall values for cutoffs used in this study are highlighted as *grey points*

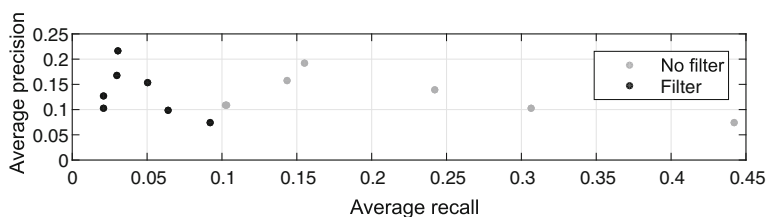
with step equal to 0.01 was evaluated with precision and with edited recall, where precision is the number of correctly predicted contacts out of all detected contacts ($precision = TP / (TP + FP)$) and $recall_{edit}$ is the number of correctly predicted contacts out of all existing contacts at the positions which were not cut out during MSA pretreatment ($recall_{edit} = (TP) / (TP + FN_{part})$). Values obtained from every protein were then averaged and plotted. Results are shown in Fig. 1, where it can be seen that $recall_{edit}$ decreases rapidly with increasing *precision*. Therefore we had to choose a compromise where *precision* and $recall_{edit}$ are in balance. This resulted in following cutoff values: $MI_{cut} = 0.3$, $RC_{cut} = 0.65$, and $JI_{cut} = 0.65$.

3.2 Comparison of Methods

The performance of every method is evaluated by precision, recall, and by improvement over a random predictor. The improvement is calculated as the ratio between precision of the method and the precision of a random predictor (N_c / N_p). Where N_c is the number of real contacts in the protein and N_p are all the possible contacts. All the efficiency evaluators are listed in Table 2, where it is also possible to

Table 2 Efficiency of different methods used for contact map prediction

Method	Precision		Recall		Improvement	
	No filter	Filter	No filter	Filter	No filter	Filter
MI	0.18	0.20	0.15	0.03	2.87	3.11
RC	0.18	0.23	0.08	0.02	2.86	3.51
JI	0.14	0.11	0.07	0.01	2.21	1.82
MR	0.15	0.18	0.20	0.04	2.48	2.88
MJ	0.12	0.12	0.18	0.04	2.02	1.97
RJ	0.13	0.14	0.14	0.03	2.09	2.25
MRJ	0.11	0.11	0.23	0.05	1.82	1.86

**Fig. 2** Average precision of each method plotted against average recall of the method to compare the performance of methods without using filters (*grey points*) and with use of filters (*black points*)

compare the efficiency of methods with and without the use of filters. It is obvious, that filtration did not bring the previously proposed benefit of precision increase without recall decrease. This can also be seen in Fig. 2, where the range of precision of methods using filters is almost the same like the precision reached by methods without filtration, furthermore the reached recall when using filters is significantly lower compared to the recall values, when filters are not used.

The performance of each new method (MR, MJ, RJ, MRJ) was first compared to the prediction performance of MI which according to Table 2 showed the best efficiency out of all single methods (MI, RC, JI). These results are shown in Fig. 3, here it can be seen that even though reached precision was lowered in majority of cases, it led to a higher recall improvement, since most of the points lie above the diagonal. By comparing Fig. 3 with results in Table 2, we can say, that MR (Fig. 3a) and MRJ (Fig. 3d) performed the best in comparison to MI. In both cases a smaller precision impairment led to a higher recall improvement. The only exception, when a new method performed worse than MI in terms of both precision and recall, occurred when RJ (Fig. 3c) was used. The use of MJ (Fig. 3b) led to a recall improvement, however not as high as the use of MR or MRJ, and for the price of worse precision impairment.

The comparison of RC analysis, which is one of the most widely used CM analysis method for protein structure prediction, with the proposed methods is visualized in Fig. 4. In this figure we can see that compared to MI comparison above, none of

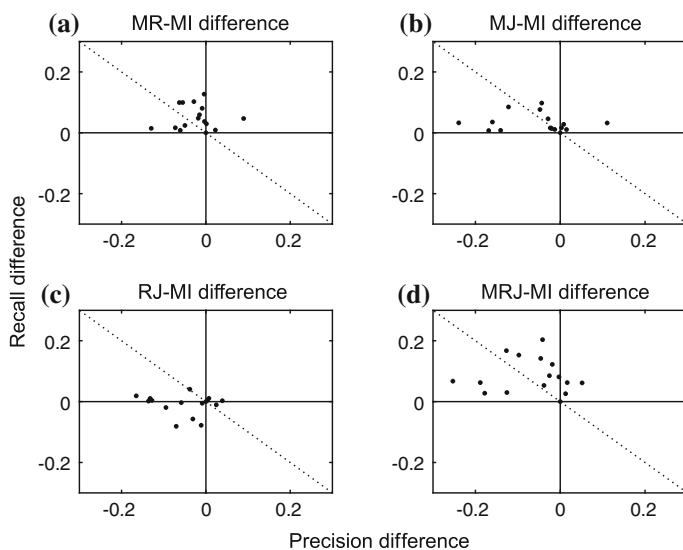


Fig. 3 Recall difference of multidimensional CM methods and MI plotted against precision difference of multidimensional CM methods and MI. Multidimensional CM methods are: **a** MR, **b** MJ, **c** RJ, and **d** MRJ. The *dashed line* represents the diagonal. The results were obtained from the whole dataset

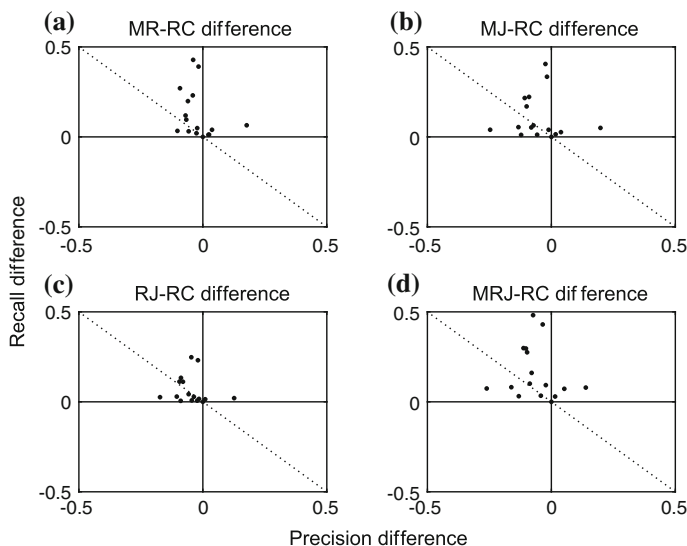


Fig. 4 Recall difference of multidimensional CM methods and RC plotted against precision difference of multidimensional CM methods and RC. Multidimensional CM methods are: **a** MR, **b** MJ, **c** RJ, and **d** MRJ. The *dashed line* represents the diagonal. The results were obtained from the whole dataset

the cases led to recall impairment. After looking at Fig. 4, we can again say that MR and MRJ showed the best results, where some cases showed almost 0.5 recall enhancement. From Table 2 we can see that 0.03 decrease in precision, when MR was used instead of RC analysis, was compensated with 0.12 rise of recall value which is more than double of the recall reached by use of only RC analysis itself. Further improvement in recall was reached when MRJ was used, however, as it can be seen in Fig. 4d, precision difference values here are more dissipated compared to precision difference values when MR is used (Fig. 4a), which leads to a higher overall precision decrease. MJ use also led to more than double increase of average recall, however, same as in MRJ, the precision difference values in Fig. 4b are more dissipated compared to the use of MR, which, in final, leads to a worse precision impairment.

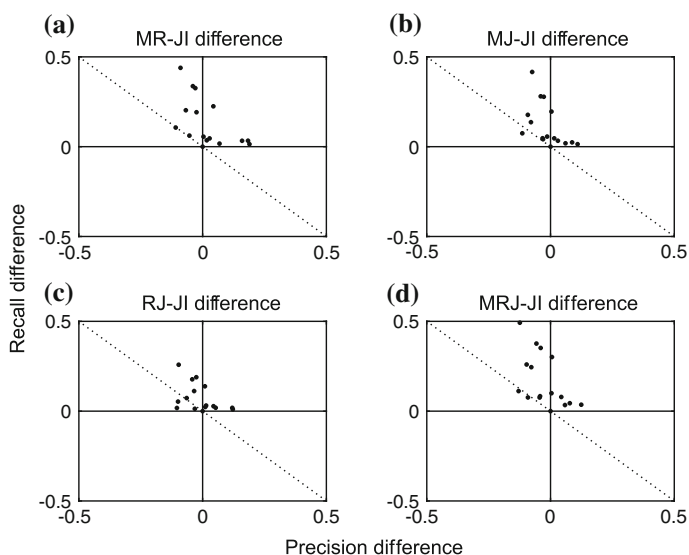


Fig. 5 Recall difference of multidimensional CM methods and JI plotted against precision difference of multidimensional CM methods and JI. Multidimensional CM methods are: **a** MR, **b** MJ, **c** RJ, and **d** MRJ. The *dashed line* represents the diagonal. The results were obtained from the whole dataset

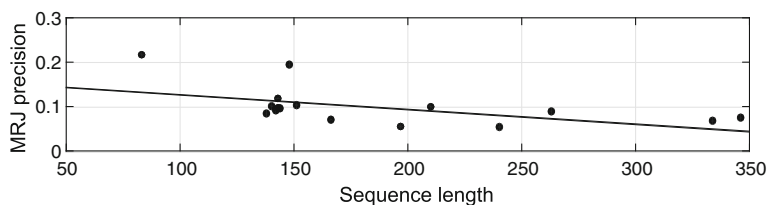


Fig. 6 Precision of MRJ plotted against protein sequence length with linear fit

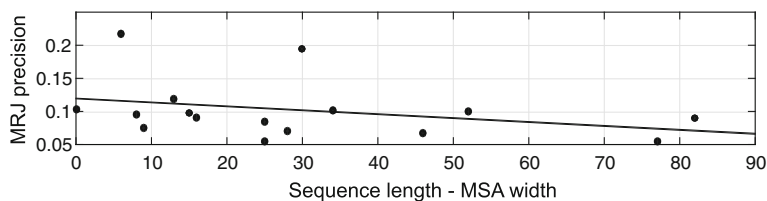


Fig. 7 Precision of MRJ plotted against the difference between sequence length and MSA width (trimming rate) with linear fit

And finally JI, which showed the worst efficiency in Table 2 when used independently, was compared to multidimensional methods with results plotted in Fig. 5. The use of all multidimensional methods led to overall increase of recall, like in RC comparison. It is possible to see in Fig. 5a, b, d that maximum recall gain reached again up to almost 0.5 which is an enormous improvement. Furthermore in about 50 % of cases an increase in precision values can be observed along with recall increase. When RJ was used, average recall value was doubled from 0.07 to 0.14 with precision loss equal to only 0.01, further, recall rise to 0.18 was gained by use of MJ with precision loss 0.02. The greatest recall growth to more than triple of original JI recall value was reached with use of MRJ with loss of only 0.03 precision. The best performance is however reached when MR is used; this leads even to precision improvement on top of almost triple recall increase.

To summarize all of the above mentioned results, we can say that MR and MRJ showed the best performance out of all four newly introduced methods. Although by use of MRJ the highest recall values are reached, precision values are slightly reduced compared to MR. Therefore, we further studied possible influence of other parameters on MRJ precision. In Fig. 6 the relationship between sequence length and reached MRJ precision is plotted. It is observed that precision lowers with increasing sequence length, thus MRJ is more suitable for shorter sequences. We also studied the impact of MSA width trimming on MRJ precision. In Fig. 7, we can see a slight trend where precision is lowered towards the higher trimming rate and therefore MRJ precision could be ameliorated by selecting slightly more diverse sequences for an MSA input.

4 Conclusion

In this study, we introduce new multidimensional approaches based on correlated mutation analysis for protein contact map prediction. Due to the use of undemanding statistical techniques providing partial 1D results which are merged into final multidimensional space, the approaches keep low computational requirements while provide substantially better results than partial techniques themselves. The outcomes of the methods were presented on several large datasets obtained from database of non-redundant protein sequences at NCBI.

Current correlated mutation analysis techniques, also those combined with other methods, e.g. neural network or other machine learning approaches, are focused on precision improvement without mentioning recall which is also important parameter. Our focus was on maximization of gained recall while minimally interfere with the other prediction efficiency parameters. With use of MR we reached 0.2 recall while maintaining the precision value at 0.15 which is impressive considering only correlated mutation analysis was used for prediction. The overall improvement of MR over a random predictor is 2.48. Using MRJ for prediction further increased recall value up to 0.23 with slightly lower reached value of precision 0.11 which is, however, still in the norm of maximum achievable precision. The improvement coefficient of MRJ is 1.82.

Even though the efficiency of statistical techniques is not sufficient enough for clear protein fold prediction, we believe that implementing MR or MRJ into a combined method, e.g. neural networks with multidimensional correlated mutations, would lead to a significant improvement in protein structure prediction that would not, otherwise, be realizable by partial techniques.

Acknowledgments Supported by the grant project GACR P102/11/1068 NanoBioTECell.

References

1. Ashkenazy, H., Unger, R., Kliger, Y.: Optimal data collection for correlated mutation analysis. *Proteins* **74**, 545–555 (2009)
2. Bouvrie, J., Baker, M.: MEA Data Analysis: Information Theoretic Approach. MIT Midterm Project Report (2004)
3. Bremm, S., Schreck, T., et al.: Computing and visually analyzing mutual information in molecular co-evolution. *BMC Bioinform.* **11**, 330 (2010)
4. Di Lena, P., Nagata, K., Baldi, P.: Deep architectures for protein contact map prediction. *Bioinformatics* **28**, 2449–2457 (2012)
5. Fariselli, P., Olmea, O., Valencia, A., Casadio, R.: Prediction of contact maps with neural networks and correlated mutations. *Protein Eng. Design Sel.* **14**, 835–843 (2001)
6. Gloor, G., Martin, L., Wahl, L., Dunn, S.: Mutual information in protein multiple sequence alignments reveals two classes of coevolving positions. *Biochemistry* **44**, 7156–7165 (2005)
7. Gobel, U., Sander, C., et al.: Correlated mutations and residue contacts in proteins. *Proteins Struct. Funct. Genet.* **18**, 309–317 (1994)
8. Gupta, N., Mangal, N., Biswas, S.: Evolution and similarity evaluation of protein structures in contact map space. *Proteins* **59**, 196–204 (2005)
9. Hamacher, K.: Relating sequence evolution of HIV1-protease to its underlying molecular mechanics. *Gene* **422**, 30–36 (2008)
10. Kundrotas, P., Alexov, E.: Predicting residue contacts using pragmatic correlated mutations method: reducing the false positives. *BMC Bioinform.* **7**, 503 (2006)
11. Li, Z., et al.: CorMut: an R/Bioconductor package for computing correlated mutations based on selection pressure. *Bioinformatics* **30**, 2073–2075 (2014)
12. Martin, L., Gloor, G., Dunn, S., Wahl, L.: Using information theory to search for co-evolving residues in proteins. *Bioinformatics* **21**, 4116–4124 (2005)
13. Olmea, O., Valencia, A.: Improving contact predictions by the combination of correlated mutations and other sources of sequence information. *Fold. Des.* **2**, 25–32 (1997)
14. Salzberg, S., Warnow, T.: Algorithms in Bioinformatics. Springer (2009)

15. Sander, C., Schneider, R.: Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins Struct. Funct. Genet.* **9**, 56–68 (1991)
16. Seemayer, S., Gruber, M., Soding, J.: CCMpred-fast and precise prediction of protein residue-residue contacts from correlated mutations. *Bioinformatics* **30**, 3128–3130 (2014)
17. Sievers, F., et al.: Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* **7**, 539 (2011)
18. Tegge, A., Wang, Z., Eickholt, J., Cheng, J.: NNcon: improved protein contact map prediction using 2D-recursive neural networks. *Nucleic Acids Res.* **37**, 515–518 (2009)
19. Xu, F., et al.: Correlated mutation analysis on the catalytic domains of Serine/Threonine protein kinases. *PLoS ONE* **4**, e5913 (2009)

Human Fibroblast Growth Factor 2 Hot Spot Analysis by Means of Time-Frequency Transforms

Anna Tamulewicz and Ewaryst Tkacz

Abstract Energy in protein complexes is not uniformly distributed. Some of amino acid residues—called hot spots—contribute most to the total energy of interaction. Hot spots can be determined experimentally or by computational methods. Here we present the application of time-frequency tools such as short-time Fourier transform and S transform to analyze human fibroblast growth factor 2 protein. The time-frequency tools take advantage the Resonant Recognition Model (RRM), which is based on the correlation between spectra of numerical representations of amino acids and their function. Thus, RRM allows for applying digital signal processing tools to amino acid analysis. Methods using time-frequency transforms do not require knowledge of protein structure, thus they help to predict hot spot residues with good accuracy and lower computational requirements comparing to other algorithms.

Keywords Hot spot · Resonant recognition model · Short-time Fourier transform · S transform · Fibroblast growth factor

1 Introduction

Most of biological processes in living cells are controlled by protein complexes. Energy on the surface of the protein complexes is not uniformly distributed. Some of amino acid residues—called hot spots—contribute most to the total energy of interaction. The term ‘hot spot’ was introduced by Clackson and Wells in 1995 in their work on the binding of human growth hormone with its receptor [2]. Hot spot residues are found to mutate slower than the rest of the protein surface, thus, hot spot is usually consisted of sequences structurally more conserved. Studies of

A. Tamulewicz (✉)

Faculty of Automatic Control, Electronics and Computer Science,
Silesian University of Technology, Gliwice, Poland
e-mail: anna.tamulewicz@polsl.pl

E. Tkacz

Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland

protein binding have shown that hot spot is surrounded by energetically less important residues causing occlusion from solvent, which is found to play an important role in highly energetic interactions. Hot spot residues are also clustered within tightly packed regions rather than randomly spread in the protein [6, 21].

Analysis of energy distribution in protein complexes is an important issue, because hot spot detection is a key to discover structures, biological function and interactions between proteins. Such studies are considerable due to possibility of controlling biological activity, which can be really useful in drugs design and curing diseases by proteins with predefined functions [21]. Another applications of hot spot identification is detection of mutations that could significantly affect the interactions in the protein complex. Also, it can be applied in searching single nucleotide polymorphisms, which may affect protein interactions [8].

1.1 Existing Methods of Hot Spot Identification

Hot spot residues can be experimentally determined by molecular biology technique called Alanine Scanning Mutagenesis. In the experiment, each amino acid of protein chain is substituted by alanine one by one. If the mutation to alanine of the analyzed amino acid residue gives the change in the binding free energy of more than 2 kcal/mol, the given residue can be deemed as hot spot [6, 13]. Hot spots from Alanine Scanning Mutagenesis experiments was deposited by Bogan and Thorn and is accessible by the Internet in Alanine Scanning Energetic Database (ASEdb) [19]. The method is widely used and it gives good results of determining hot spot residues, however it is time-consuming and expensive, which induces the need of developing other methods e.g. computational.

The one of the pioneering work of hot spot prediction is algorithm Robetta, which was introduced in 2002 by Kortemme and Baker [7]. The basis of the method is free energy function, which takes into account Lennard-Jones interactions, Coulomb electrostatics, solvation interactions and hydrogen bonding. Thus, the model includes shape complementarity, polar interactions (ion pairs and hydrogen bonds), interactions between protein atoms and the solvent. The algorithm uses an atomic representation of the protein. It determines hot spot residues by predicting the change in the binding free energy upon alanine mutation basing on the free energy function, therefore the method is also called computational alanine scanning [7, 8].

Another algorithm representing different approach to the problem of detecting hot spots is a method called HotPoint developed in 2010 by Tuncbag et al. [20, 21] and it is used in a HotPoint server available online. The base of the method is empirical model, which takes into account conservation, solvent accessibility, statistical pairwise potentials of residues and change in the binding free energy.

Other computational method worth mentioning is MAPPIS (Multiple Alignment of Protein-Protein Interfaces) [14], which detects spatially conserved interaction patterns by performing multiple alignments of the physico-chemical interactions and the

binding properties in three dimensional space. It takes into account spatial conservation of interactions formed between groups of atoms in protein-protein interfaces in order to predict hot spots. Grosdidier et al. [4] presented the docking approach in the method predicting hot spots by applying normalized interface propensity (NIP) values obtained from rigid-body docking simulations without prior knowledge of the complex structure. The model considers hot spots contributing to the interaction due to electrostatic and water-to-interface desolvation effect.

The methods mentioned above have one significant drawback—they require knowledge of protein or complex structure, however structure has been solved only for small percentage of known proteins. Despite the fact that biological functions of protein are related to its structure, it is assumed that all information about protein structure and functions is embedded in its primary structure (amino acid sequence). Thus, it can be possible to analyze hot spots using only amino acid sequence of protein [9]. In recent years it became more popular to use this assumption in order to search for hot spots employing digital signal processing methods [5, 11–13].

2 Resonant Recognition Model

The basic idea of using digital signal processing approach to analysis of protein hot spots is provided by so called Resonant Recognition Model (RRM) [3], which assumes correlation among function of protein and numerical representation of its amino acids. The RRM is a physico-mathematical approach that interprets protein sequence using signal analysis methods. Algorithms basing on Resonant Recognition Model tend to have lower computational requirements than methods involving knowledge of protein structure.

2.1 *Conversion of Amino Acids Sequences into Numerical Signal*

In order to use any digital signal processing method to amino acid analysis, it is necessary to convert amino acid sequence into numerical signal basing on some physical and biochemical properties relevant for protein's biological activity. Studies have shown [3, 22] that electron-ion interaction pseudo-potential (EIIP) is a good descriptor of energy distribution in protein, thus, it can represent a physical characteristic of amino acids and correlates with some of their biological properties. EIIP describes average energy state of delocalized electrons of the amino acid, which is critical for protein biological function (i.e. interaction with protein's target). Numerical signal can be obtained by assigning each amino acid specific EIIP value according to Table 1.

Table 1 EIIP values of amino acids

Amino acid	3 letter code	1 letter code	EIIP
Alanine	Ala	A	0.0373
Aspartic acid	Asp	D	0.1263
Phenylalanine	Phe	F	0.0946
Histidine	His	H	0.0242
Lysine	Lys	K	0.0371
Methionine	Met	M	0.0823
Proline	Pro	P	0.0198
Arginine	Arg	R	0.0959
Threonine	Thr	T	0.0941
Tryptophan	Trp	W	0.0548
Cysteine	Cys	C	0.0829
Glutamic acid	Glu	E	0.0058
Glycine	Gly	G	0.0050
Isoleucine	Ile	I	0.0000
Leucine	Leu	L	0.0000
Asparagine	Asn	N	0.0036
Glutamine	Gln	Q	0.0761
Serine	Ser	S	0.0829
Valine	Val	V	0.0057
Tyrosine	Tyr	Y	0.0516

2.2 Characteristic Frequency

According to the RRM, there is a correlation between the amplitude spectrum of numerical representation of amino acids and their biological function. Veljkovic et al. [22] observed that functionally related proteins share a spectral component in Fourier transforms of their numerical sequences. The frequency of this component is called characteristic frequency of the functional group and is obtained by multiplying spectrum of numerical representation of analyzed protein with spectra of its related proteins (i.e. sharing the same function, but found in different organism). The product function of this multiplication is called consensus spectrum. There is a significant peak (characteristic frequency) in the consensus spectrum of functionally related proteins. In the case of functionally unrelated proteins such peaks have not been found.

The characteristic frequency of K biologically related proteins can be calculated by finding the frequency of significant peak in the consensus spectrum, which is determined by computing the cross-spectral function $M(n)$ described by (1):

$$M(n) = |X_1(n)| |X_2(n)| \dots |X_K(n)|, \quad (1)$$

where X_1, X_2, \dots, X_K are the Discrete Fourier Transforms corresponding to K proteins. One of hot spots feature is structural conservation, which allows for using RRM to hot spot analysis.

3 Time-Frequency Analysis

Most of biomedical signals—such as amino acid sequences—are non-stationary, i.e. their characteristics change with time. This feature makes the standard frequency tools insufficient for analysis of such signals. Fourier transform can be useful to search for specific frequency, but it is not able to find the time of its occurrence and duration. Time-frequency analysis comprises the information about frequency components and time intervals of their occurrence in non-stationary signals [10]. In this section we describe the most popular and basic transform for time-frequency analysis—short-time Fourier transform—and its generalization—S transform.

3.1 The Short-Time Fourier Transform

The short-time Fourier transform (STFT) is Fourier-related transform, which is useful for analyzing non-stationary signals. It requires multiplication of the analyzed signal with time window, then the Fourier transform of the shortened signal is computed. The time window is shifted by predefined period of time and the procedure of computing Fourier transform is repeated [10].

STFT provides time-frequency representation of the signal:

$$STFT(\tau, f) = \int_{-\infty}^{\infty} h(t)g(\tau - t)e^{-i2\pi ft} dt, \quad (2)$$

where $h(t)$ denotes signal to be transformed, τ —the time of spectral localization, f —the Fourier frequency, $g(t)$ —a window function and i —imaginary unit [23].

The biggest drawback of STFT is its fixed resolution. It is important to choose proper window width, because too narrow window gives better time resolution but poor frequency resolution (a general Heisenberg cube properties are valid). On the other hand, too wide window gives better frequency resolution but poor time resolution [10]. Because of fixed window, STFT cannot track dynamics of non-stationary signal properly. It is impossible to provide good frequency and time resolution simultaneously, although there is a guideline how to choose optimal window [24]. Thus, sometimes it is difficult to select the best window width for specific signal analysis.

3.2 The S Transform

The drawback of STFT described above can be partially overcome by continuous wavelet transform (CWT), although CWT itself also has some disadvantages. In CWT, sinusoidal basis functions of Fourier transform and windows of STFT are replaced by scaled versions of a mother wavelet. As a result, CWT contains information similar to STFT, but it also provides a progressive resolution, which is a trade-off between time and frequency resolution. The drawback of CWT is that it does not measure frequency directly, because instead of complex sinusoidal basis functions, it uses rather a related concept, called scale [1].

The S transform (ST) was developed in 1996 by Stockwell et al. [15] for analyzing geophysics data. It can be treated as an extension of CWT with phase correction or STFT with variable window (Gaussian window function). ST combines properties of both CWT and STFT and overcomes their common drawbacks [23].

The one of the methods of achieving the ST is to derive it from STFT (2) by replacing the window function $g(t)$ with the Gaussian function:

$$g(t) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}}. \quad (3)$$

Thus, the ST can be defined:

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-2i\pi ft} dt. \quad (4)$$

The S transform of discrete time series $h(kT)$ with a sampling interval of T and $k = 0, 1, \dots, N - 1$ can be obtained by letting $t \rightarrow kT$, $f \rightarrow n/NT$, $\tau \rightarrow jT$ and is defined as:

$$S\left(jT, \frac{n}{NT}\right) = \sum_{m=0}^{N-1} H\left(\frac{m+n}{NT}\right) e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{2\pi imj}{N}}, \quad (5)$$

where $j, m, n = 0, \dots, N - 1$ and H denote discrete Fourier transform:

$$H\left(\frac{n}{NT}\right) = \frac{1}{N} \sum_{k=0}^{N-1} h(kT) e^{-\frac{2\pi ink}{N}}. \quad (6)$$

3.3 Time-Frequency Filtering

Both the STFT and S transform are invertible, i.e. original signal can be recovered from the transforms by the inverse transforms, which allows for time-frequency filtering.

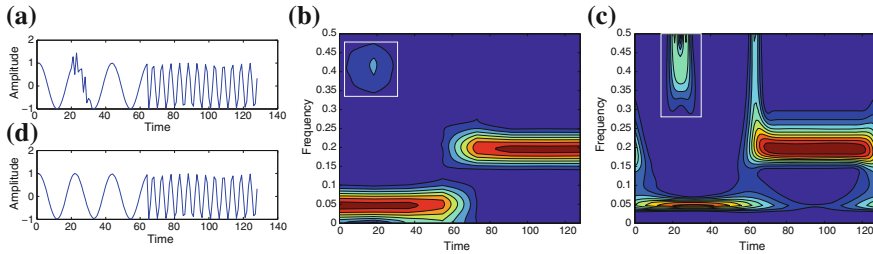


Fig. 1 **a** The synthetic time series consisting of low-frequency signal (time: 0–63), middle-frequency signal (time: 64–128) and high-frequency noise (time: 20–30); **b** STFT spectrum of the synthetic time series. The high frequency noise is marked by *white rectangle*; **c** ST spectrum of the synthetic time series. The high frequency noise is marked by *white rectangle*; **d** The denoised signal recovered by removing noise from ST spectrum and applying inverse ST

The inverse of the STFT can be defined by:

$$h_{inv}(\tau) = \frac{1}{g(0)} \int_{-\infty}^{\infty} STFT(\tau, f) e^{i2\pi f\tau} df, \quad (7)$$

whereas the discrete inverse ST is shown in:

$$h_{inv}(kT) = \sum_{n=0}^{N-1} \left\{ \frac{1}{N} \sum_{j=0}^{N-1} S\left(jT, \frac{n}{NT}\right) \right\} e^{\frac{i2\pi nk}{N}}. \quad (8)$$

Figure 1 presents an example of discrete time series, its spectra using STFT and ST and signal recovered using inverse transform. Note that ST provides better frequency resolution for low-frequency signal and better time resolution for high-frequency noise than STFT, while STFT provides better frequency resolution for high-frequency signal.

In order to detect hot spots using time-frequency filtering we used algorithm similar to ones described in [11, 13]:

1. Converting analyzed amino acid sequence and its related protein into numerical signal using EIIP values.
2. Computation of consensus spectrum and characteristic frequency.
3. Computation of the spectra (ST and STFT) of the analyzed amino acid sequence.
4. Generating the filtering matrix selecting the characteristic frequency from the spectra (ST and STFT).
5. Filtering of the amino acid numerical sequence by using the filtering matrix.
6. Finding the localization of hot spots by locating the energy peaks in the filtered signal:

$$E(kT) = |h_{inv}(kT)|^2. \quad (9)$$

4 Human Fibroblast Growth Factor 2 Analysis

4.1 Data Set

For the purpose of this paper we chose human fibroblast growth factor 2 (FGF2) to present the application of signal processing methods to the hot spots analysis. FGF2 is a protein from fibroblast growth factor family, which plays an important role in the regulation of cell survival, cell division, angiogenesis, cell differentiation and cell migration [18].

Primary sequences of analyzed proteins were acquired from freely accessible UniProt database [16, 17]. Only the extent of polypeptide chains in the mature proteins following processing were analyzed. We used two data sets; first of them (data set A) comprises related proteins of human FGF2 described in [13] and is presented in Table 2; in the second data set (data set B) related proteins were obtained by finding FGF2 from different organisms (Table 3).

Table 2 Data set A

Swiss-Prot ID	Protein name	Organism
P05230	FGF1	Homo sapiens
P15656	FGF5	Mus musculus
P55075	FGF8	Homo sapiens
O15520	FGF10	Homo sapiens
O54769	FGF16	Rattus norvegicus
Q9EPC2	FGF23	Mus musculus
Q9HCT0	FGF22	Homo sapiens
Q9QY10	FGFP1	Rattus norvegicus
P03969	FGF2	Bos taurus

Related proteins of human fibroblast growth factor 2. Data set contain FGF protein from different organisms

Table 3 Data set B

Swiss-Prot ID	Protein name	Organism
P12226	FGF2	Xenopus laevis
Q5IS69	FGF2	Pan troglodytes
P20003	FGF2	Ovis aries
P48798	FGF2	Monodelphis domestica
P13109	FGF2	Rattus norvegicus
P15655	FGF2	Mus musculus
P48800	FGF2	Gallus gallus
P03969	FGF2	Bos taurus

Related proteins of human fibroblast growth factor 2. Data set contain FGF2 protein from different organisms

4.2 Time-Frequency Analysis

The analysis of human FGF2 hot spots was conducted according to the algorithm described in Sect. 3.3. The results of FGF2 analysis is presented in Figs. 2, 3, 4, 5, and 6. Figure 2 shows the consensus spectrum of both analyzed data sets. Note the additional peaks in consensus spectrum of data set B, which—according to [3, 22]—can suggest that the group of proteins has more than one common function. For data set A only one significant peak was found which imply that this data set is more suitable for hot spot analysis.

Figure 3 presents amplitude of spectra (STFT and ST) obtained from numerical signal of analyzed amino acid sequence (human fibroblast growth factor 2). The characteristic frequency calculated for data set A is marked by line. The difference

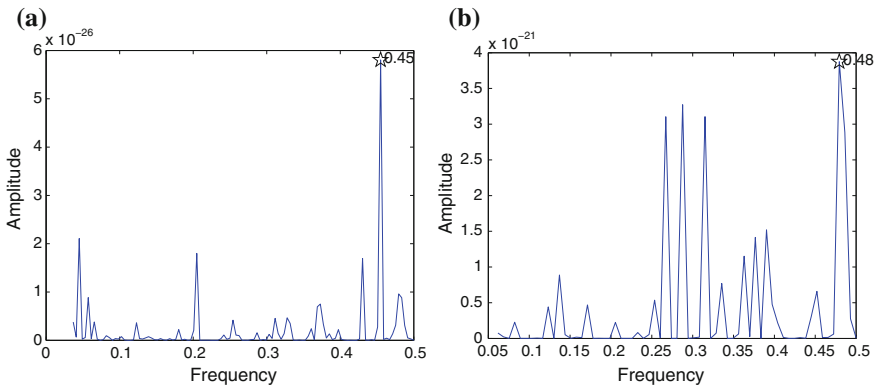


Fig. 2 Consensus spectrum of human FGF2 functionally related proteins. The peak in the consensus spectrum correspond to the characteristic frequency and is marked by a star. **a** Consensus spectrum of data set A. **b** Consensus spectrum of data set B

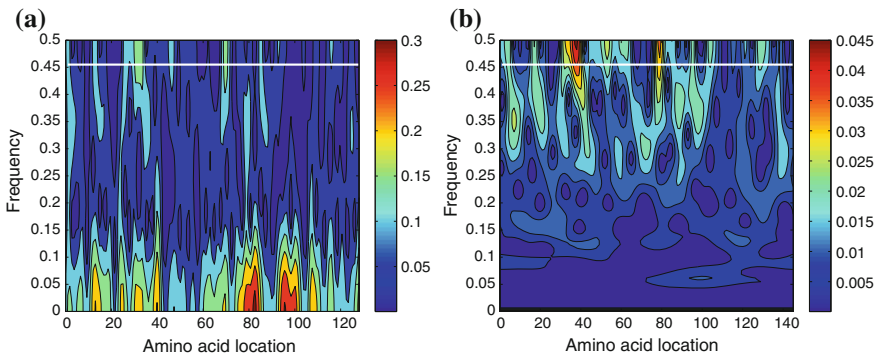


Fig. 3 **a** STFT spectrum of the human FGF2; **b** ST spectrum of the human FGF2. The characteristic frequencies of data set A is marked by line

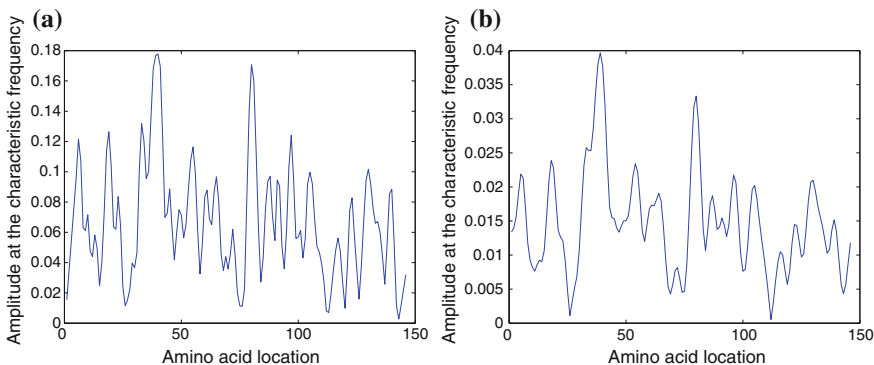


Fig. 4 **a** Slice of STFT spectrum at characteristic frequencies (0.45) for data set A; **b** Slice of ST spectrum at characteristic frequencies (0.45) for data set A

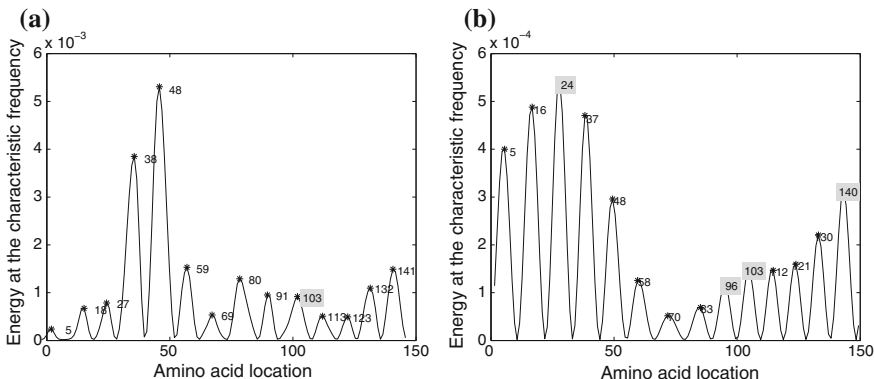


Fig. 5 Energy at the characteristic frequency (0.45) using data set A. Hot spot locations covering hot spots from ASEdb are marked by shade. **a** Energy obtained from STFT spectrum; **b** Energy obtained from ST spectrum

between STFT spectrum and ST spectrum can be observed in the slices of the spectra at the characteristic frequency (Fig. 4).

In order to detect hot spot residues, the energy at the characteristic frequency was calculated according to (9). The peaks in the energy plots correspond to calculated hot spot locations. Figures 5 and 6 present energy at the characteristic frequency for data set A and B respectively, obtained both from STFT and ST spectra. Thus, 4 sets of hot spots were received, using: data set A and STFT (Fig. 5a), data set A and ST (Fig. 5b), data set B and STFT (Fig. 6a), data set B and ST (Fig. 6b). Results were compared with ASEdb, which contains hot spots from experimental technique (ASM). According to ASEdb, locations of hot spot amino acids in human FGF2 protein is: 24, 96, 103 and 140. Hot spots found using ST and data set A are: 5, 16, 24, 37, 48, 58, 70, 83, 96, 103, 112, 121, 130, 140. Note that all hot spots described in

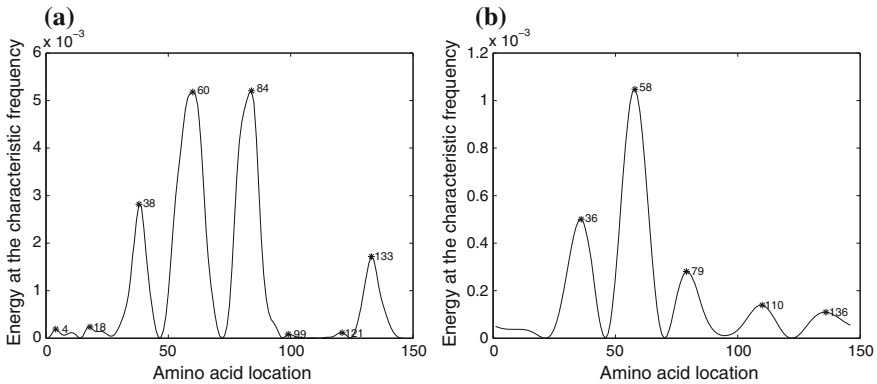


Fig. 6 Energy at the characteristic frequency (0.48) using data set B. None of the identified hot spot locations cover hot spots from ASEdb. **a** Energy obtained from STFT spectrum; **b** Energy obtained from ST spectrum

ASEdb were discovered by applied algorithm and additional 10 were found, which can suggest that there might be more hot spots in the FGF2 protein, which would require further experimental validation. The worse result was obtained using STFT and data set A; only one of hot spot location comparing to ASEdb was found correctly. Although, in the energy plot, there are also peaks close to location of hot spots from ASEdb (hot spot locations: 27 vs 24, 91 vs 96, 141 vs 140). We noticed that the results obtained using STFT and ST often differ by few amino acid locations.

Entirely different results was obtained for data set B—none of the identified hot spot locations overlap with hot spots form ASEdb, using both the STFT and ST, which confirms that the data set is not a good choice for hot spot analysis, even though the characteristic frequency is not significantly different (0.03) from characteristic frequency of data set A.

5 Summary and Conclusion

Hot spots identification in protein complexes may help to understand interactions between proteins. Experimental techniques require a lot of effort, but computational methods are suitable for determining hot spots from non-hot spots and may help reveal unknown hot spots, which will need further experimental validation. Standard models are able to predict energetically important amino acid residues with satisfying results comparing to experimental data, but they usually require the knowledge of protein complex and higher computational requirements comparing to digital signal processing methods.

In this paper, we analyzed amino acid sequence of human fibroblast growth factor 2 using signal processing methods employing time-frequency tools such as short-time

Fourier transform and S transform in order to present application of different tools for hot spot analysis and compare results obtained by these methods. The introduced algorithms base on Resonant Recognition Model, which allows for applying digital signal processing tools to amino acid analysis. The analyzed amino acid sequences was converted into numerical signal by assigning each amino acid specific EIIP value. We used two data sets of related proteins and calculated consensus spectrum to find characteristic frequencies. In order to determine which amino acids are responsible for performance on characteristic frequency, spectra (short-time Fourier transform and S transform) of human FGF2 were obtained. Then, the characteristic frequency was filtered from the spectra and energy was calculated, which allowed to find locations of hot spot amino acids.

We used two different data sets to present the importance of the selection of the related proteins set and its impact on results. The best results for human FGF2 sequence was obtained using ST and data set A, which allowed for finding all hot spots comparing to ASEdb and 10 newly identified. Thus, in the set of the 14 identified hot spots, 4 are true positive and other 10 are false positive comparing to the experimental data. Although, the false positive hot spots are often interpreted as unknown hot spots, not yet identified by experiments and thus requiring further validation. STFT revealed worse results—only 1 of detected hot spots overlaps with ASEdb. We also found that selection of related proteins is very important. Even though the difference between characteristic frequencies for both analyzed data sets was subtle, the results vary significantly. For second of analyzed data set, the detected hot spots do not overlap with hot spots from ASEdb. The consensus spectrum for this data set suggests that the set of related proteins is not suitable for the analysis, because it revealed more than one significant peak.

The paper showed that it is possible to combine amino acid analysis with digital signal processing with quite satisfying results.

Acknowledgments This work was supported by the Ministry of Science and Higher Education funding for statutory activities of young researchers (decision no. 8686/E-367/M/2015 of 12 March 2015).

References

1. Brown, R.A., Frayne, R.: A fast discrete S-transform for biomedical signal processing. In: Conference Proceedings IEEE Engineering Medicine Biology Society, vol. 2008, pp. 2586–2589 (2008)
2. Clackson, T., Wells, J.A.: A hot spot of binding energy in a hormone-receptor interface. *Science* **267**(5196), 383–386 (1995)
3. Cosic, I.: Macromolecular bioactivity: is it resonant interaction between macromolecules?—theory and applications. *IEEE Trans. Biomed. Eng.* **41**(12), 1101–1114 (1994)
4. Grosdidier, S., Fernandez-Recio, J.: Identification of hot-spot residues in protein-protein interactions by computational docking. *BMC Bioinform.* **9**, 447 (2008)
5. Kasperek, J., Maderankova, D., Tkacz, E.: Protein hotspot prediction using S-transform. *Adv. Intell. Syst. Comput.* **283**, 327–336 (2014)

6. Keskin, O., Ma, B., Nussinov, R.: Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* **345**(5), 1281–1294. Epub 2004 Dec 2 (2005)
7. Kortemme, T., Baker, D.: A simple physical model for binding energy hot spots in protein-protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **99**(22), 14116–14121 (2002)
8. Kortemme, T., Kim, D.E., Baker, D.: Computational alanine scanning of protein-protein interfaces. *Sci. STKE* **2004**(219), pl2 (2004)
9. Ofran, Y., Rost, B.: Protein-protein interaction hotspots carved into sequences. *PLoS Comput. Biol.* **3**(7), e119 (2007)
10. Pięciak, T., Rzeszutko, J.: Wprowadzenie do analizy i przetwarzania sygnałów biomedycznych. in: R. Tadeusiewicz, P. Augustyniak, *Podstawy inżynierii biomedycznej*, tom I, Wydawnictwo AGH, Kraków (2009) (in Polish)
11. Ramachandran, P., Antoniou, A.: Identification of hot-spot locations in proteins using digital filters. *IEEE J. Sel. Top. Signal Process.* **2**(3), 378–389 (2008)
12. Ramachandran, P., Antoniou, A., Vaidyanathan, P.P.: Identification and location of hot spots in proteins using the short-time discrete Fourier transform. In: *Proceedings 38th Asilomar Conference Signals, Systems, Computers*, Pacific Grove, CA, Nov 2004, pp. 1656–1660 (2004)
13. Sahu, S.S., Panda, G.: Efficient localization of hot spots in proteins using a novel S-transform based filtering approach. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **8**(5), 1235–1246 (2011)
14. Shulman-Peleg, A., Shatsky, M., Nussinov, R., Wolfson, H.J.: Spatial chemical conservation of hot spot interactions in protein-protein complexes. *BMC Biol.* **5**, 43 (2007)
15. Stockwell, R.G., Mansinha, L., Lowe, R.P.: Localization of the complex spectrum: the S transform. *IEEE Trans. Signal Process.* **44**(4), 998–1001 (1996)
16. The UniProt Consortium: UniProt: a hub for protein information. *Nucleic Acids Res.* **43**(Database issue), D204–D212 (2015)
17. The Universal Protein Resource (UniProt), European Bioinformatics Institute (EMBL-EBI), Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR), <http://www.uniprot.org> (2015)
18. The Universal Protein Resource (UniProt), European Bioinformatics Institute (EMBL-EBI), Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR), <http://www.uniprot.org/uniprot/P09038> (2015)
19. Thorn, K.S., Bogan, A.A.: ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* **17**(3), 284–285 (2001)
20. Tuncbag, N., Keskin, O., Gursoy, A.: HotPoint: hot spot prediction server for protein interfaces. *Nucleic Acids Res.* **38**(Web Server issue), W402–W406 (2010)
21. Tuncbag, N., Gursoy, A., Keskin, O.: Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* **25**(12), 1513–1520 (2009)
22. Veljković, V., Cosić, I., Dimitrijević, B., Lalović, D.: Is it possible to analyze DNA and protein sequences by the methods of digital signal processing? *IEEE Trans. Biomed. Eng.* **32**(5), 337–341 (1985)
23. Wang, Y.-H.: The Tutorial: S-Transform. <http://djj.ee.ntu.edu.tw/Stransform.pdf> (2010)
24. Zieliński, T.P.: *Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań*, Wydawnictwa Komunikacji i Łączności, Warszawa (2005) (in Polish)

Structured Bi-clusters Algorithm for Classification of DNA Microarray Data

Pawel Foszner and Andrzej Polanski

Abstract We propose a new algorithm for classification of transcriptomic data based on the two stage procedure of feature selection. The construction of the new feature set is based on the hypothesis that in many transcriptomic datasets there is an additional hidden structure dictated by some biological factors, which were not taken into account in the design of the experiment. The hidden structure in the data is detected by using the specialized version of the bi-clustering methodology, called the method of structured bi-clusters. The idea of the new method is that the constructed bi-clusters must coincide along one dimension with the existing data classes. The second dimension of the bi-clusters is a free design parameter of our algorithm. Combining the proposed feature selection procedure with the nearest neighbour classification rule leads to improvements of classification accuracy for real transcriptomic data.

Keywords Bi-clustering · Feature selection · Machine learning · Data mining

1 Introduction

Technologies of DNA microarrays based measurement of transcriptome profiles in cells allowed for obtaining large volumes of data on biological and cellular processes and mechanisms behind all aspects of functioning of organisms and cells, such as mechanisms of growth and development, disease versus normal states, responses to environmental conditions and many others [1, 12, 14, 17]. Application of these data resulted in significant development of knowledge in biology and medicine. Many of the created DNA microarray datasets are continuously available in publicly open repositories and can be repetitively applied when verifying and refining detailed aspects of interpretation of transcriptome profiles. One of the routine approaches to

P. Foszner (✉) · A. Polanski
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Informatics, Silesian University of Technology, Gliwice, Poland
e-mail: Pawel.Foszner@polsl.pl
URL: <http://www.polsl.pl>

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_14

the analysis of DNA microarray data is by conducting computational experiments involving construction/design of molecular classifiers and estimation of classification accuracy achieved by these classifiers in random validation procedures [2, 5]. Such computational experiments allow for estimating the nature and quality of the biological information in the collected dataset—by examining classification accuracy, and for studying biological mechanisms behind studied phenomena by examining lists of gene probes chosen as gene signatures in the feature selection steps in the classifier design procedures. By studying literature references on DNA data classification one can easily see that virtually all classification algorithms were already applied and compared in problems of DNA data analysis, e.g., linear discrimination rules [11], SVD [23], neural networks [3], nearest neighbours [24], random forests [7], nearest centroids [20], and many others. At present there is rather little or no progress in DNA classification related to improvement of classification of DNA microarray data in the sense of the choice of the classification rules. However, one field of the research is still under intensive development. It concerns procedures of feature selection. Due to the large numbers of probes in DNA chips and high levels of noise, existing procedures of feature selection can be unstable and prone to many errors. Some of these errors are related to multiple testing when constructing gene signatures. Other important sources of errors are related to strong correlations between expression profiles of genes belonging to the same gene pathways. There are, therefore, many efforts focused on improving procedures for feature selection in the design of molecular classifiers in the aspects of their quality and stability [8, 9, 15]. It has already been demonstrated that improving the procedure of feature selection in the molecular classifier can lead to both better classification accuracy and to deeper insights to biological background behind the performed experiments. In this paper we present a new algorithm for improving classification accuracy of molecular classifiers by proposing a new idea of the feature selection. The hypothesis behind our construction of feature selection procedure is that in many transcriptomic datasets there is an additional hidden structure in the data related to some biological factors, which were ignored in the design of the experiment. Existence of such factors is probable due to the complicated nature of the design of biological experiments. We propose the algorithm, called structured bi-clusters, for the detection of such hidden structure in the datasets, and for feature selection based on the detected hidden structure. The hidden structure in the data is detected by using the idea of bi-clustering [4, 16, 19]. A specialized version of bi-clustering methodology was developed, such that the constructed bi-clusters must coincide along one dimension with the existing data classes. The second dimension of the bi-clusters is the free construction parameter of our algorithm. We combine the proposed the feature selection procedure with the nearest neighbour classification rule and we apply the obtained molecular classifier to several datasets for which quality of classification was already extensively studied in the literature [20]. We demonstrate improvements achieved thanks to our design.

2 Description of the Structured Bi-clusters Algorithm

Suppose that we have a large number of microarray experiments conducted on a group of patients with a different type of cancer. Gene expression array will consist of as many rows as we have genes and so many columns as we have patients. If we use this matrix as the input to bi-clustering algorithm, we expect to receive bi-clusters consisting of sub-groups of genes having similar expression profile for the subgroup of patients with a certain type of cancer each. With so defined the problem we are dealing with unsupervised algorithm, where each group will be discovered as latent information. But many times in the case of a microarray (for other applications as well), only groups in one dimensions are hidden information. Groups of second dimension are known in advance. In the case of microarray experiments, we know exactly the groups of patients, but we cannot directly answer the question of which groups of genes relate to them. Bi-clustering certainly solves this problem, but unsupervised approach to the subject results in a significant loss of information right from the start. In this paper, we propose modifications of bi-clustering approaches based on matrix decomposition algorithms. The modifications consist in the introduction to the process of detecting groups of knowledge of one of them. These changes have made the process of detecting supervised (from the point of view of one of the groups).

2.1 Notation

Notation was taken from the paper by Madeira and Oliveira [18], where bi-cluster is defined by a subset of rows and subset of columns from data matrix. Given the data matrix V with set of rows (X), and set of columns (Y), a bi-cluster (B) is defined by a sub-matrix (I, J), where I is a subset of X , and J is a subset of Y .

$$A = (X, Y), \tag{1}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{bmatrix}, X_i = [a_{i1} \ a_{i2} \ \cdots \ a_{iN}], Y = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{Mj} \end{bmatrix}, \tag{2}$$

$$B = (I, J), I \in X, J \in Y. \tag{3}$$

Single bi-clustering experiment (R) outputs K bi-clusters, where K is a number which, depending on the algorithm used, can be a parameter given by the user, or the number formed as a result of executing the selected method.

$$R = \{B_i = (I_i, J_i)\}, \text{ where } i = 1, \dots, K \text{ and } \forall : I_i \in X, J_i \in Y. \tag{4}$$

2.2 Non-negative Matrix Factorization

A very wide range of algorithms are algorithms based on data matrix decomposition. In such methods data matrix (A) is factorized into (usually) much smaller matrices. Such a distribution, because of the much smaller matrices is much easier to analyse, and the obtained matrices reveal previously hidden features. These algorithms are often called NMF algorithms. NMF stands for non-negative matrix factorization. Two efficient algorithms were introduced by Seung and Lee [16]. First minimize conventional least square error distance function and second generalized Kullback-Leibler divergence. Third and last from this group is algorithm that slightly modify the second approach. Author [21] introduce smoothing matrix for achieving a high degree of sparseness, and better interpretability of the results. Data matrix in this techniques is factorized into (usually) two smaller matrices:

$$A \approx WH. \quad (5)$$

Finding the exact solution is computationally very difficult task. Instead, the existing solutions focus on finding local extrema of the function describing the fit of the model to the data. AspectAnalyzer implements five algorithms based on non-negative matrix factorization:

- PLSA witch stands for Probabilistic Latent Semantic Analysis. Introduced by Thomas Hoffman [13], and based on maximizing log-likelihood function. For this purpose author use Expectation-Maximization algorithm [6].
- Based on minimization of Least Square Error distance function

$$\|A - WH\|^2 = \sum_{ij} (A_{ij} - WH_{ij})^2. \quad (6)$$

- Based on minimization of Kullback-Leibler divergence

$$D(A || WH) = \sum_{ij} (A_{ij} \log \frac{A_{ij}}{WH_{ij}} - A_{ij} + WH_{ij}). \quad (7)$$

- Based on minimization of non-smooth Kullback-Leibler divergence

$$D(A || WSH) = \sum_{ij} (A_{ij} \log \frac{A_{ij}}{WSH_{ij}} - A_{ij} + WSH_{ij}), \quad (8)$$

$$S = (1 - \theta)\mathbf{I} + \frac{\theta}{q}\mathbf{1}\mathbf{1}^T. \quad (9)$$

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 2 & 3 & 4 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 4 & 6 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 6 & 9 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 8 & 12 & 16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 1 Bi-cluster extraction in NMF algorithms

2.3 Supervised Modification

In non-negative matrix factorization algorithms single bi-cluster is obtain using single row vector from matrix W and single (corresponding) column vector from matrix H . Multiplication of those two vectors represent data matrix related only with one selected bi-cluster (Fig. 1).

If data matrix $A^{[p,n]}$ consist of p rows and n columns than factorized matrices are $W^{[p,k]}$ and $H^{[k,n]}$. Common dimension k is usually provided by the user as an input parameter. Bi-clusters are extracted as follows:

$$\begin{aligned}
 B_i &= (I_i, J_i) \quad \text{where } i = 1, \dots, K \\
 I_i &= \{x_l \in X : \forall l x_l > TI_i\}, \\
 J_i &= \{y_l \in Y : \forall l y_l > TJ_i\}.
 \end{aligned}
 \tag{10}$$

Thresholds TI_i and TJ_i in case of example from Fig. 1 can be set to 0. In real life applications due to noise in input data and in resulted bi-clusters, those values are usually set to vector mean or higher. The higher the values for these thresholds, the smaller bi-cluster (but with better quality). Setting them depends on the user and his needs.

In case of gene expression analysis only one group of attributes are hidden (usually its group of genes) and group of samples are known in advanced. It is good to introduce this knowledge to analysis in order to improve its quality. All NMF algorithms at its start, form matrices W and H using random values, and after some series of update steps in this matrices bi-clusters are formed. In our experiments we modify vectors of W or H (depends of this what dimension in known) in such a way that we arbitrary decides about which group will be associated with the vector.

Modification is possible due to fact that in bi-clusters $B_i = (I_i, J_i)$ vectors I_i or J_i is known in advanced. After initializing matrices W and H its values are modified in order to introduce this knowledge. Each vector I_i is connected with one row from matrix W (or each vector J_i with one column from matrix H). Then all attributes that are present in vector I_i (or J_i) are double in i 'th row of matrix W (or i 'th column in matrix H). All other attributes are reduced twice. With such initialized data bi-clusters are formed according to our alignment and its quality is significantly better. Thus the groups associated with the second dimension (the one that is not known) is more accurate and better discriminate the collection.

3 Data

Real data set was taken from publication of Ronsenwald et al. [22]. Data consist of 240 patients from witch were taken tissues for microarray analysis, each of witch were measured 7399 genes. Samples were divided into three different subgroups. Authors decided to assign 160 samples into training set and 80 into validation set, and we follow this division.

The authors in theirs paper proved that the presented data are not easy to classify. They have achieved the percentage of correctly classified samples at 65%. Three year later Stefan Michiels et al. [20] tried to improve this result. They use nearest centroid approach to filter the attributes and at the end improve classification. After theirs filtration was left 6693 genes, and no significant gain in classification rate was achieved.

4 Results

All results and input data are available free to download at http://foszner.pl/pub/supplementary/struct_biclusters/. On the site you will find the following files:

- train.arff—The original set of training data,
- validation.arff—The original set of validation data,
- result_training{ID}.arff—The original set of training data with a reduced set of attributes. {ID} represents the experiment ID,
- result_validation{ID}.arff—The original set of validation data with a reduced set of attributes. {ID} represents the experiment ID.

All bi-clustering experiments were performed using AspectAnalyzer software package [10] available at <http://aspectanalyzer.foszner.pl> as ready to use installer with extensive user manual. AspectAnalyzer generated a ARFF WEKA files with filtered attributes. The classification was realized using the WEKA software version 3.6. Due to the fact that all used software is free and all data (input and output) have been provided as supplementary materials—these results are easily verifiable.

Table 1 PLSA algorithm

ID	Class. Rate (%)	Attrib. Num.	Gain (%)
5891	75	368	7.5
5892	78.75	273	11.25
5893	76.25	723	8.75
5894	76.25	443	8.75
5898	70	514	2.5
5899	70	245	2.5
5901	75	438	7.5
5902	71.25	365	3.75
5903	75	475	7.5
5933	71.25	2193	3.75
5934	65	469	-2.5
5935	70	888	2.5
5937	70	417	2.5
5938	61.25	1230	-6.25
5939	78.75	458	11.25
5940	72.5	627	5
5941	61.25	475	-6.25
5942	73.75	687	6.25
5943	71.25	836	3.75
5944	71.25	311	3.75

Bi-clustering analysis were performed only on training data set. After the successful analysis filtered subset of genes were used to create training and validation data set. Training file were used to train classifier and it was tested on validation file only. As classifier was used the nearest neighbours algorithm with the following parameters:

```
weka.classifiers.lazy.IBk -K 3 -W 0 -A
"weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -
R first-last\""
```

The original data without the filtered parameters obtained classification rate at 67.5%. Tables 1, 2, 3 and 4 represent the various experiments that resulted in data with filtered attributes. The comparison provided the experiment ID, the percentage of correctly classified samples in validation set, and comparison between accuracy of classification with and without our structured bi-clusters filtering stage. The gain parameter reported in the last column of Tables 1, 2, 3 and 4 is defined as the difference between numbers of correctly classified samples in the validation set. Positive sign means that the version of classification algorithm with structured bi-clusters step improves classification, while negative signs depict the opposite situation.

Table 2 Least square error algorithm

ID	Class. Rate (%)	Attrib. Num.	Gain (%)
5896	68.75	1243	1.25
5914	75	796	7.5
5915	77.5	259	10
5916	72.5	2009	5
5917	70	821	2.5
5918	62.5	261	-5
5919	80	357	12.5
5920	83.75	546	16.25
5921	68.75	1379	1.25
5922	73.75	584	6.25
5955	77.5	494	10
5956	73.75	483	6.25
5957	72.5	1425	5
5958	72.5	427	5
5959	68.75	457	1.25
5960	75	526	7.5
5961	75	554	7.5
5962	66.25	1224	-1.25
5963	82.5	438	15
5964	67.5	521	0

Table 3 Kullback-Liebler algorithm

ID	Class. Rate (%)	Attrib. Num.	Gain (%)
5904	67.5	777	0
5905	77.5	1108	10
5906	71.25	1411	3.75
5907	67.5	1493	0
5908	65	1249	-2.5
5909	76.25	427	8.75
5910	62.5	2750	-5
5911	71.25	1362	3.75
5912	77.5	485	10
5913	71.25	471	3.75
5945	81.25	374	13.75
5946	71.25	5479	3.75
5947	81.25	539	13.75
5948	73.75	1270	6.25
5949	78.75	412	11.25
5950	72.5	567	5
5951	78.75	711	11.25
5952	67.5	4692	0
5953	78.75	778	11.25
5954	75	1591	7.5

Table 4 Non-smooth Kullback-Liebler

ID	Class. Rate (%)	Attrib. Num.	Gain (%)
5923	72.5	346	5
5924	75	181	7.5
5925	78.75	388	11.25
5926	78.75	282	11.25
5927	67.5	248	0
5928	83.75	469	16.25
5929	72.5	183	5
5930	76.25	199	8.75
5931	81.25	230	13.75
5932	76.25	288	8.75
5965	82.5	266	15
5966	78.75	322	11.25
5967	86.25	153	18.75
5968	77.5	467	10
5969	71.25	240	3.75
5970	66.25	166	-1.25
5971	81.25	254	13.75
5972	78.75	157	11.25
5973	87.5	249	20
5974	80	241	12.5

Table 5 Summary of fitting results into normal distribution

Method	Num. of Res.	Mean (%)	Std. (%)	Max value (%)
PLSA	100	71.09	5.75	81.25
K-L	100	72.06	5.73	83.75
LSE	100	72.6	6.17	83.75
nsK-L	100	74.71	6.16	90

The tables mentioned above only contain a randomly selected results for each of the selected methods. All conducted experiments are summarized in Table 5. The results (classification rates) for each method has been adjusted to a normal distribution. Table present mean values and standard deviations for the modeled distributions, and in addition—maximum values for each group of results.

5 Conclusion

We have analysed real microarray data set including 240 cancer patients, divided into three groups. We have implemented three-class classification and we performed training validation experiments with the same structure of training (160 patients) and validation (80 patients) sets as that reported in the paper [22]. The classical version of the multi-class classifier was nearest neighbours classifier, well suited to multi-class tasks, with numbers of neighbours optimized for the analysed dataset.

We have compared two variants of classification algorithms, classical nearest neighbours as described in the previous paragraph, and nearest neighbours preceded by detection of structured bi-clusters, as described in the Methods section. As we show in Tables 1, 2, 3 and 4 we achieve, on the average, significant improvement of classification accuracies.

Acknowledgments The work was performed using the infrastructure supported by POIG.02.03.01-24-099/13 grant: “GeCONiI—Upper Silesian Center for Computational Science and Engineering”.

References

1. Afshari, C.A., Nuwaysir, E.F., Barrett, J.C.: Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res.* **59**(19), 4759–4760 (1999)
2. Ancona, N., Maglietta, R., Piepoli, A., D’Addabbo, A., Cotugno, R., Savino, M., Liuni, S., Carella, M., Pesole, G., Perri, F.: On the statistical assessment of classifiers using DNA microarray data. *BMC Bioinform.* **7**(1), 387 (2006)
3. Arena, P., Bucolo, M., Fortuna, L., Occhipinti, L.: Cellular neural networks for real-time DNA microarray analysis. *IEEE Eng. Med. Biol. Mag.* **21**(2), 17–25 (2002)
4. Cheng, Y., Church, G.M.: Biclustering of expression data. In: *Ismb*, vol. 8, pp. 93–103 (2000)
5. Cho, S.B., Won, H.H.: Machine learning in dna microarray analysis for cancer classification. In: *Proceedings of the First Asia-Pacific Bioinformatics Conference on Bioinformatics 2003*, vol. 19, pp. 189–198. Australian Computer Society, Inc. (2003)
6. Dempster, A.P., Laird, N.R.D.: Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. Ser. B* **39** (1977)
7. Díaz-Uriarte, R., De Andres, S.A.: Gene selection and classification of microarray data using random forest. *BMC Bioinform.* **7**(1), 3 (2006)
8. Ding, C., Peng, H.: Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **3**(02), 185–205 (2005)
9. Dрамиński, M., Rada-Iglesias, A., Enroth, S., Wadelius, C., Koronacki, J., Komorowski, J.: Monte carlo feature selection for supervised classification. *Bioinformatics* **24**(1), 110–117 (2008)
10. Foszner, P., Polanski, A.: Aspectanalyzer—distributed system for bi-clustering analysis. In: *Man-Machine Interactions 4, Advances in Intelligent Systems and Computing*, vol. 391, pp. 395–404 (2016)
11. Guo, Y., Hastie, T., Tibshirani, R.: Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* **8**(1), 86–100 (2007)
12. Heller, M.J.: DNA microarray technology: devices, systems, and applications. *Annu. Rev. Biomed. Eng.* **4**(1), 129–153 (2002)

13. Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. *Mach. Learn. J.* 177–196 (2001)
14. Khan, J., Bittner, M.L., Chen, Y., Meltzer, P.S., Trent, J.M.: DNA microarray technology: the anticipated impact on the study of human disease. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1423**(2), M17–M28 (1999)
15. Koronacki, J., Cwik, J.: *Statistical Learning Systems*. WNT, Warsaw (in Polish) (2005)
16. Lee, D., Seung, S.: Algorithms for non-negative matrix factorization. *Advances in Neural Information Processing Systems*, pp. 556–562 (2000)
17. Lettieri, T.: Recent applications of dna microarray technology to toxicology and ecotoxicology. *Environ. Health Perspect.* 4–9 (2006)
18. Madeira, S.C., Oliveira, A.L.: Biclustering algorithms for biological data analysis: a survey. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 24–45 (2004)
19. Michalak, M., Lachor, M., Polański, A.: Hibi—the algorithm of biclustering the discrete data. In: *Artificial Intelligence and Soft Computing*, pp. 760–771. Springer (2014)
20. Michiels, S., Koscielny, S., Hill, C.: Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* **365**(9458), 488–492 (2005)
21. Pascual-Montano, A., Carazo, J.M., Kochi, K., Lehmann, D., Pascual-Marqui, R.D.: Non-smooth nonnegative matrix factorization (nsNMF). *IEEE Trans. Pattern Anal. Mach. Intell.* 403–415 (2006)
22. Rosenwald, A., Wright, G., Chan, W.C., Connors, J.M., Campo, E., Fisher, R.I., Gascoyne, R.D., Muller-Hermelink, H.K., Smeland, E.B., Giltman, J.M., Hurt, E.M., Zhao, H., Averett, L., Yang, L., Wilson, W.H., Jaffe, E.S., Simon, R., Klausner, R.D., Powell, J., Duffey, P.L., Longo, D.L., Greiner, T.C., Weisenburger, D.D., Sanger, W.G., Dave, B.J., Lynch, J.C., Vose, J., Armitage, J.O., Montserrat, E., López-Guillermo, A., Grogan, T.M., Miller, T.P., LeBlanc, M., Ott, G., Kvaloy, S., Delabie, J., Holte, H., Krajci, P., Stokke, T., Staudt, L.M.: The use of molecular profiling to predict survival after chemotherapy for diffuse large-b-cell lymphoma. *N. Engl. J. Med.* **346**(25), 1937–1947 (2002)
23. Simek, K., Fujarewicz, K., Świerniak, A., Kimmel, M., Jarzqb, B., Wiench, M., Rzeszowska, J.: Using SVD and SVM methods for selection, classification, clustering and modeling of DNA microarray data. *Eng. Appl. Artif. Intell.* **17**(4), 417–427 (2004)
24. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**(6), 520–525 (2001)

Prediction of Newborn Weight Using Questionnaire Data and Machine Learning Approach

Justyna Pieter, Sebastian Student, Karolina Sobczyk
and Krzysztof Fujarewicz

Abstract Low birth weight is an important factor of the health of an infant because of being a major determinant of morbidity, disability, and mortality during infancy and childhood. Low birth weight is a multifaceted public health problem. The prevalence of this diagnosis is estimated to be 15–20 % worldwide (more than 20 million infants) and occurs mostly in developing countries (95.6 %). There are many factors, which affect the birth weight of newborns. From a medical point of view, it is important to find out, which factors are important and determine the risk of not normal birth weight. In this work, we propose to use the classification system to find out which features are important to classify the neonates into birth weight groups. In presented approach, we use different classification and feature selection methods to analyze a new questionnaire-based obstetric data set. Our results show that a small number of features is enough to achieve high classification accuracy rate and the set of selected features is significant from biological and medical point of view.

Keywords LDA · DLDA · QDA · LBW · Obstetric data · Machine learning

1 Introduction

Low birth weight is an important factor of a health of an infant because of being a major determinant of morbidity, disability, and mortality during infancy and childhood. Furthermore, low birth weight also has a long-term impact on health outcomes in an adult life. A baby's low weight at birth is either the result of birth before 37 weeks of gestation or due to intrauterine growth restriction (IUGR) defined as a

J. Pieter (✉) · S. Student · K. Fujarewicz
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Automatic Control, Silesian University of Technology,
Akademicka 16, Gliwice, Poland
e-mail: justyna.pieter@polsl.pl

K. Sobczyk
School of Public Health, Department of Health Economics and Management,
Medical University of Silesia, Piekarska 18, Bytom, Poland

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_15

condition where a baby's growth slows or ceases when it is in the uterus. World Health Organization (WHO) defined birth weight as the first weight of the fetus or a newborn obtained after birth. For live births, birth weight should preferably be measured within the first hour of life. Low birth weight (LBW) is the birth weight lower than 2.500 g (up to and including 2.499 g). Infants born weighing less than 1.500 g are classified as a very low birth weight (VLBW) and those with a birth weight less than 1000 g—as an extremely low birth weight (ELBW). Low birth weight is a multifaceted public health problem. The prevalence of this diagnosis is estimated to be 15–20 % worldwide (more than 20 million infants) and occurs mostly in developing countries (95.6 %). In Organization for Economic Co-operation and Development (OECD) countries, one-in-fifteen babies are born with a weight lesser than 2500 g (6.7 % of all births). In Poland, it is about 6.1 %. The prevalence of low birth weight infants has increased in most OECD countries since 1980. Among the reasons for this growth are the increasing number of multiple births, older age at childbearing and increased in the use of induction of labor and caesarean delivery, which have increased the survival rates of low birth weight babies. Due to the fact that low birth weight is a global problem WHO's the Member States have endorsed global target for low birth weight policy brief, which goal is to achieve a 30 % reduction in the number of infants born with a low birth weight by the year 2025 (reduction from approximately 20 million to about 14 million infants with low weight at birth). There are many factors, which affect the birth weight of newborns. They may be related to the mother, the infant, or the physical environment and play an important role in determining not only the birth weight but also health condition in the future. Among the maternal risk factors for low birth weight maternal age can be mentioned (adolescent motherhood or pregnancy over 35 years of age), a previous history of low weight births and miscarriages, having poor nutrition, a low body mass index (BMI), stress level during pregnancy and lifestyle factors such as excessive alcohol consumption and smoking (including exposure to second-hand smoke). Moreover, one should not forget about the socio-economic factors, like educational level, family income, housing conditions and mother's employment status.

2 A Questionnaire Data Set Description

A cross-sectional hospital-based and questionnaire-based study was conducted in Clinical Ward of Obstetrics and Gynecology in Ruda Slaska, Medical University of Silesia, Poland from April 15th, 2013 to January 12th, 2014. The examined group consisted of 1355 women hospitalized in connection with childbirth and their newborns.

In hospital-based study the research material was obtained from the medical records of the following types:

- hospital patient questionnaire (medical history),
- obstetrical card (the course of hospitalization, obstetric history, the overall condition of the fetus, parturition, health status of the newborn, features afterbirth),
- delivery observation card (the course of labor),
- postnatal development history (the course of labor, health status of the newborn at birth and during hospitalization, observation, and medical orders).

The criterion for inclusion in the study was a hospital stay associated with a delivery-pregnant solution at 23 weeks of gestation or later (ICD-10 codes between O80–O84). Exclusion criteria was a hospitalization in connection to pregnancy with the abortive outcome—pregnancy outcomes before 23 weeks of gestation (ICD-10 codes between O00–O08) or multiple gestations as a cause of hospitalization (ICD-10 code O30). In the questionnaire-based study, the instrument for data collection was constructed based on the literature and the main analyze was preceded by a pilot study that led to refining the questionnaire. The final version of the questionnaire consist of four sections which include:

- socio-demographic and economic data (age, marital status, education, profession, place of residence, housing conditions, income),
- maternal risk factors in pregnancy (active and passive smoking, alcohol consumption, BMI, physical activity, dietary supplementation, stress),
- prenatal care (a type of insemination, gynecological-obstetric care, prenatal genetic testing),
- childbirth (a type of delivery, newborn sex, birth weight, birth length).

The criterion for inclusion in this part of the study was an outcome of delivery as single live birth. Exclusion criteria was an outcome of delivery as multiple birth or stillbirth or neonatal death prior to the second/third day after birth—before attending the patient in the questionnaire-based study.

3 Methodology

Optimal design of the classification system for public health medical phenological data described in the previous section constitutes the challenging, computational and conceptional problem. To the most important and crucial steps of classifier design we include data preprocessing, choice of proper classification and feature selection methods. For that reason, we compare results from different classification, selection methods and a different number of selected features. Our classifier is based on the classification system proposed in [6] and is shown in Fig. 1.

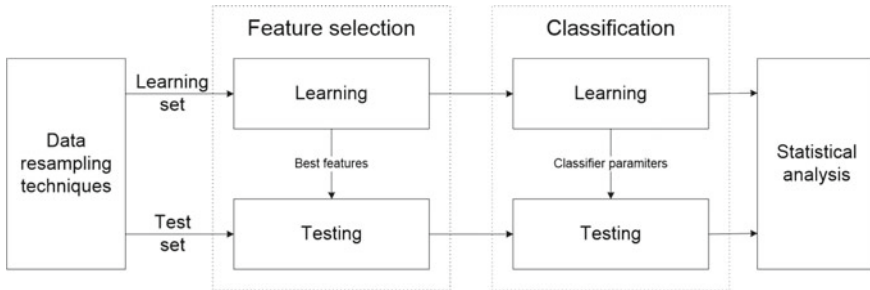


Fig. 1 Classification system scheme

3.1 Data Preprocessing

First of all, we needed to make a preprocessing. We wanted to find the features that could be the most important, and which we suspected to have the precious information. We removed from the collected data set features which obviously influence the birth weight. In the next step all patients with non-complete data was removed. Thanks to such treatment we obtained a data array that can be analyzed by supervised machine learning algorithms.

Normally the birth weight data can be grouped using the International Classification of Diseases (ICD-10) in the part concerning disorders of newborn related to the short gestation and low birth weight and another related to long gestation and high birth weight. To classify the group of neonates first was used the following ICD-10 codes:

- P07.0—extremely low birth weight newborn (≤ 1000 g),
- P07.1—other low birth weight newborn (1000–2499 g),
- P08.0—exceptionally large newborn baby (2500–4499 g),
- P08.1—other newborn heavy- or large-for-dates regardless of period of gestation (4000–4499 g).

All the other newborns with a birth weight in the range 2500–3999 g were assigned to subgroups named *normal birth weight*.

For classification purposes we used two-class prediction model of the birth weight with two class: *over 2500* and *below 2500* which is slightly different that five classes used in the International Classification of Diseases (ICD-10). Because the birth weight depends on the pregnancy week (HBD) we removed this bias by additional correction.

3.2 Feature Selection and Classification Methods

We have compared the following methods of classification: Linear Discriminant Analysis classifier (LDA), Diagonal Linear Discriminant Analysis classifier (DLDA) [4, 5] and Quadratic Discriminant Analysis (QDA). Additionally, we used following feature selection algorithms: a stable selection method (gs) proposed in [7], the BSS/WSS-statistics, where BSS denotes the between-group sum of squares and WSS the within-group sum of squares (BSS_WSS) [3], the modified partial least squares implementation SIMPLS (SIMPLS) [2] and the classical t-statistics (t-test).

3.3 LDA and DLDA Methods

The second classification method that we have used is Linear Discriminant Analysis. The primary purpose of LDA is to separate samples of distinct K groups. The k th sample group, $k \in \{1, 2, \dots, K\}$, has a class mean

$$\bar{x}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_{k,i}, \quad (1)$$

where N_k is a number of observations in k th class and $N = \sum_{k=1}^K N_k$ is the total number of observations.

Let us define a sample group covariances matrix

$$\sigma_k = \frac{1}{N_k - 1} \sum_{i=1}^{N_k} (x_{k,i} - \bar{x}_k)(x_{k,i} - \bar{x}_k)^T, \quad (2)$$

a scatter matrix between class

$$S_b = \sum_{k=1}^K N_k (\bar{x}_k - \bar{x})(\bar{x}_k - \bar{x})^T \quad (3)$$

and a within class covariance matrix

$$S_w = \sum_{k=1}^K (N_k - 1) \sigma_k \quad (4)$$

computed by pooling the estimates of the covariance matrices of each class.

The main objective of LDA is to find a projection matrix Φ_{LDA} that maximizes the ratio of the determinant of S_b to the determinant of S_w .

$$\Phi_{LDA} = \arg \max_{\Phi} \frac{||\Phi^T S_b \Phi||}{||\Phi^T S_w \Phi||}, \quad (5)$$

which may be found by solving the following equation

$$S_b \Phi - S_w \Phi \Lambda = 0. \quad (6)$$

In DLDA, the covariance matrix σ_k is estimated by the diagonal common sample covariance matrix. The covariance matrix is diagonal with each diagonal element being the pooled sample variance of the corresponding predictor.

3.4 QDA Methods

The quadratic discriminant analysis is closely related to the linear discriminant analysis. In QDA, it is assumed that the measurements from each class are normally distributed and we assume that covariances matrix can be different for each class. Because of that we need to estimate the covariances matrix σ_k for each k th class.

Quadratic discriminant function for k th class is defined as follows:

$$\delta_k(x) = -\frac{1}{2} \log |\sigma_k| - \frac{1}{2} (x - \bar{x}_k)^T \sigma_k^{-1} (x - \bar{x}_k) + \log \frac{N_k}{N}. \quad (7)$$

This function is similar to the linear discriminant function except for the fact that covariance matrices σ_k are not identical and the function contains second order terms.

The classification is based on the rule

$$\hat{G}(x) = \arg \max_k \delta_k(x). \quad (8)$$

QDA, because it allows for more flexibility for the covariance matrix, tends to fit the data better than LDA, but then it has more parameters to estimate.

4 Results

Figure 2 presents classification accuracy of DLDA classifier. This figure and all next figures presented in this section show how the accuracy depends on increased number of used features. For DLDA classifier the best results was observed for the BSS WSS selection method.

Figure 3 presents classification accuracy of LDA classifier. The results was almost the same as when we have used DLDA. So the best accuracy was achieved for the BSS WSS selection method.

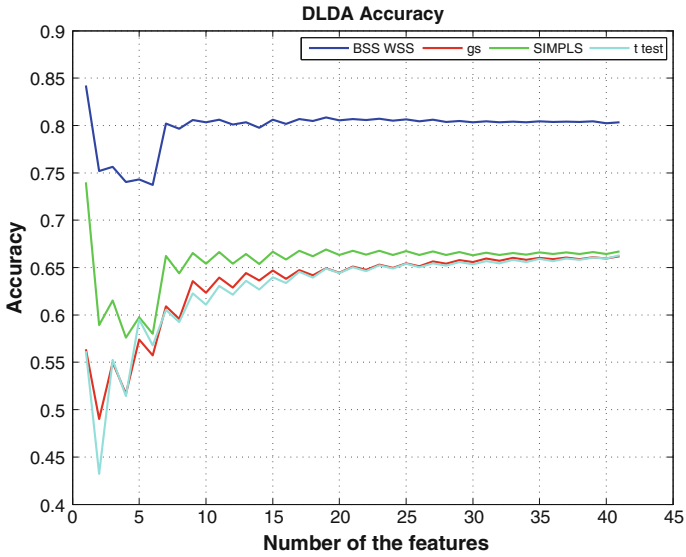


Fig. 2 Bootstrap based classification accuracy of the DLDA classifier for different number of features and different selection methods

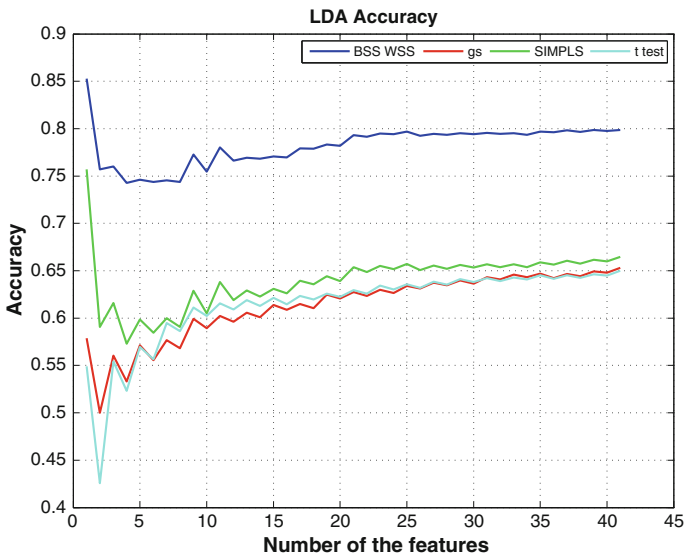


Fig. 3 Bootstrap based classification accuracy of the LDA classifier for different number of features and different selection methods

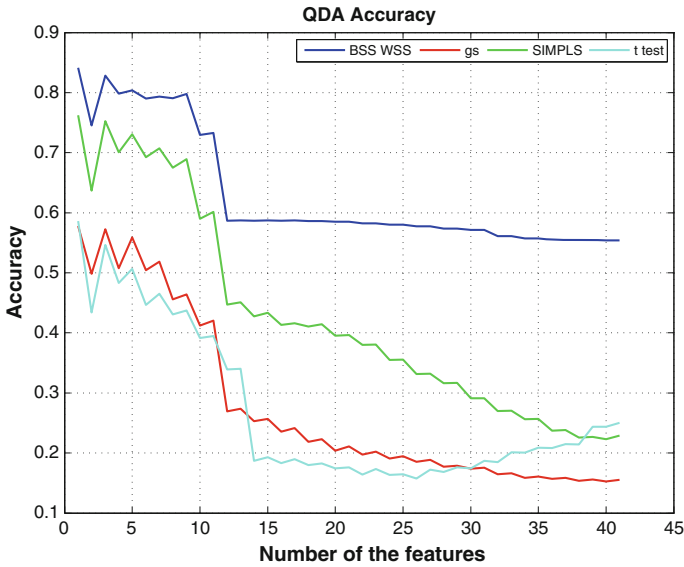


Fig. 4 Bootstrap based classification accuracy of the QDA classifier for different number of features and different selection methods

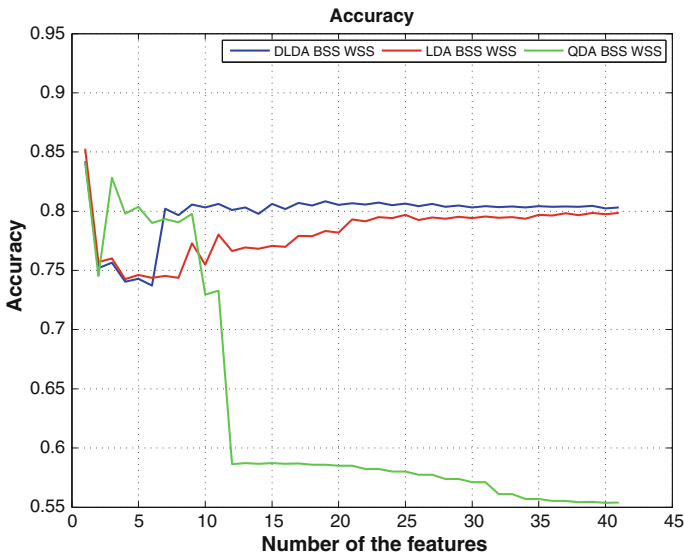


Fig. 5 Bootstrap based classification accuracy of all analyzed classifiers for different number of features

Table 1 Table of features after classification for analyzed baby’s data

Symbol	Description	Correlation	p_value
PHACT BEFPR	Physical activity before pregnancy	–	0.002
IRON SUPP	Iron supplementation	–	0.002
SMOKE ALLPR	Smoking during all pregnancy	–	0.005
FOLIC SUPP	Folic supplementation	–	0.227
SMOKE ITR	Smoking in the first trimester of pregnancy	–	<0.001
EDUC F	Father’s education	+	0.011
NO SMOKE	No smoking in pregnancy	+	<0.001
SMOKE PASS	Passive smoking in pregnancy	+	<0.001
EDUC M	Mother’s education	+	0.006
MAGN SUPP	Magnesium supplementation	–	0.002

Figure 4 presents classification accuracy of QDA classifier. The best results was observed for the BSS WSS selection method.

Figure 5 presents classification accuracy of all analyzed classifiers with the best selection method.

The best result for a small number of features was received from QDA classification with BSS WSS selection. When we used more features the better accuracy was achieved by DLDA classification. Basing on the estimated train error for training data we selected best features, ranked as shown in Table 1.

5 Conclusion

The correct data preprocessing turns out to be very important and necessary stage of the analysis. The results of the described data analysis allowed us to find a couple of interesting associations between analysed attributes and the weight of the newborn. These attributes include factors such as physical activity before pregnancy, supplementation of selected minerals and vitamins, as well as parents’ educational level and pregnant mother’s exposure to passive smoking. The result may indicate that inactive women may give birth to infants with lower birth weight and have decreased risk for delivering an infant with macrosomia. Furthermore, the results suggest that iron and magnesium supplementation during pregnancy is correlated with infant birth weight. Similar observations have been shown for supplementation of folic acid. The data also shed some addition light on the relationship between parents’ educational level and birth weight. For children of mothers and fathers with higher education reported a lower risk of low birth weight. Many studies have documented a strong association between active smoking during pregnancy with fetal growth retardation. The results of this analysis also suggest a link between passive smoking and the risk of newborn

low birth weight. Because of the existence of many risk factors for low birth weight, further research on the subject at a sufficiently large representative populations is recommended.

Acknowledgments This work was supported by the Institute of Automatic Control under Grant No. BKM-514/RAU1/2015/ t.10 (JP), Grant No. BK-227/RAU1/2015/ t.3 (KF) and NCBI Grant No. PBS3-B3-32-2015 (SS).

References

1. Boser, B.E., Guyon, I.M., Vapnik, V.: A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory, ACM, pp. 144–152 (1996)
2. De Jong, S.: An alternative approach to partial least squares regression. *Chemometr. Intell. Lab. Syst.* **18**, 25–263 (1993)
3. Dudoit, S., Fridlyand, J., Speed, P.: Comparison of discrimination methods for classification of tumors using gene expression data. *J. Am. Stat. Assoc.* **97**, 77–87 (2002)
4. Huang, D., Quan, Y., He, M., Zhou, B.: Comparison of Linear Discriminant Analysis Methods for the Classification of Cancer Based on Gene Expression Data. Oxford University Press (1988)
5. Krzanowski, W.J.: Principles of multivariate analysis: a user's perspective. *J. Exp. Clin. Cancer Res.* **28**(1), 149 (2009)
6. Student, S., Fajarewicz, K.: Stable feature selection and classification algorithms for multiclass microarray data. *Biol. Direct* **7**(33), 1–20 (2012)
7. Yang, K., Cai, Z., Li, J., Lin, G.: A stable gene selection in microarray data analysis. *J. BMC Bioinform.* **7**, 228 (2006)

Texture Analysis to Trophoblast and Villi Detection in Placenta Histological Images

Zaneta Swiderska-Chadaj, Tomasz Markiewicz, Robert Koktysz
and Wojciech Kozlowski

Abstract The paper presents two methods of an automatic tropho-blasts and villi detection in histological images to support the pathomorphological diagnostic procedure. The studied slides represent the placenta villi from spontaneous miscarriage stained with the Hematoxylin and Eosin. The proposed methods are based on texture analyses, as Local Binary Pattern and Unser, and mathematical morphology operations. The research on placenta villi detection and the evaluation on the histological images is needed to support clinical studies. The results of the automatic trophoblasts and villi detection were compared with the expert's annotations. The average coverage of the detected trophoblast areas is 93.65 % for the Unser- method and 77.06 % for the LBP method. The obtained results confirm efficiency of the proposed solutions.

Keywords Local binary pattern · Unser · Placenta villi · Histopathological images · Texture analysis

1 Introduction

Computational pathology is a rapidly developing area. Studying histopathological images of tissue is a popular method of tissue examination, and it allows for the formulation of digital algorithms for tissue analysis. This method can be used as a support to the traditional examinations, and for the definition of prognostic indicators. The placenta is an organ which connects the developing fetus to the uterine wall to allow nutrient uptake. Trophoblasts are specialized cells of the placenta, which provide nutrients to the embryo and develop into a large part of the placenta. They have an important role in the embryo implantation and the interaction with

Z. Swiderska-Chadaj (✉) · T. Markiewicz
Warsaw University of Technology, 1 Politechniki Sq., Warsaw, Poland
e-mail: zaneta.swiderska@gmail.com

T. Markiewicz · R. Koktysz · W. Kozlowski
Military Institute of Medicine, 128 Szaserow Str., Warsaw, Poland

the decidualised maternal uterus. Placenta villi provide maximum contact area with maternal blood. Abnormal placental villi are larger and have more central blood vessels. As a result, the diffusion distance for gas and nutrient exchange is larger and, therefore, placental function is impaired. Villous changes often are associated with various fetal severe lesions even intrauterine death [10].

Histological evaluation of aborted placenta may be very important in the cases of repeated miscarriages, and for prediction and explanation of further misconception, and family planning which are one of the most significant problem in contemporary population. The most common placental pathologies are: placental infarction (25 % of cases), chorioamnionitis (12.5 % of cases), molar changes (9 % of cases), hypervascularisation, avascularisation etc. [4]. The observed indicators are the effect of perinatal pathology. The pathologic examination of the placenta provides valuable information relating to the placenta and fetal pathologies. Histological evaluation may be also very important in forensic pathology studies. This study can confirm clinical diagnoses, and provide definitive diagnosis from the clinical differential diagnosis lists [3]. For example, avascularisation indicates a problem with angiogenesis of the fetal, hypervascularisation can be an effect of disorders in the osmotic transfer between the fetal and the mother, edema is connected with the hydropic problem, and changes in circumference trophoblast proliferation are related with molar changes. Moreover, hemorrhages in villi suggest coagulation disorders. The specific examination and identification of the pathologies mentioned above is a crucial step for a future care of a patient in pre pregnancy and pregnancy period in order to avoid the next miscarriage. The cytostatics, steroids or vitamin balance treatment play a crucial role in this matter. In order to indicated properly the depicted histological placenta pathologies, the trophoblast and villi should be identified in a specimen. Moreover, most of mentioned indicators are presented in quantitative form. Thus, to develop and increase objectivity of the results, an automatic evaluation is indicated.

The morphological evaluation of the specimen requires mainly an identification of placental villi. There is no existing method to detect placental structures automatically because of the large variation in the shape, colour, and texture. Very scarce number of papers devoted to the recognition of some objects existing in the placenta have been reported. The paper [1] presents a method to automatically detect and extract blood vessels from a given histological placenta image by using image processing techniques and neural networks. Another paper [2] presents the automatic recognition of four grades of the placental tissues development. It was done on the basis of ultrasound images, which represent different ways of tissue imaging. Their approach was based on the selection of tissues, feature extraction by discrete wavelet transform and classification by the multilayer perceptron. The limitation of this approach is that the ultrasound image does not allow for the identification of the specific objects forming placenta. No works presenting the automatic recognition of placenta structures have been reported up to now. The research on placenta villi detection and evaluation on the pathomorphological images is still needed to support clinical studies.

In this work, we propose two new methods for the detection of a trophoblast that creates villi walls in the pathomorphological images stained by Hematoxylin and Eosin (H&E). They are based on the algorithm containing Local Binary Pattern (LBP) and Unser texture analysis. The quantification results of an automatic detection of trophoblast are compared with the standard visual selection performed by the expert.

2 Materials

The twelve cases of placenta villi from spontaneous miscarriage subject to H&E staining were obtained from the archives of the Department of Pathomorphology from the Military Institute of Medicine in Warsaw, Poland. The data collection was approved by the IRB of the Military Institute of Medicine. The acquisition of WSI (whole slide images) was performed on the 3DHistech Panoramic II scanner. The images were acquired under magnification $200\times$ with a resolution $0.389\ \mu\text{m}$ per pixel, saved as RGB image.

3 Methods

In this paper, we propose two different methods for the detection of villi surrounding trophoblast, based on LBP [5–8], and Unser [12] features. The proposed methods were applied to different colour representations. The computational approach is validated by the quantitative comparison with the expert results.

3.1 Colour Representations

In this paper, we propose a texture descriptors method for trophoblast detection. Colour representation is an important factor in the texture analysis. In various colour representations, different cell structures are visible to greater or smaller extend. The challenge is to choose the colour components for the best differentiation of the trophoblast. RGB is the original colour representation of the analyzed images. It is desirable to propose such a colour representation of an image in which texture descriptors used allow for differentiation of trophoblast from the tissue area. As a result, we have conducted the research for different colour representations, such as: RGB, CMYK, Luv, HSV, YCbCr. Furthermore, we have suggested to apply SCC (Special Colour Component) colour representation which consists of an average of sum of G and B colour components from RGB, in accordance with:

$$SCC = (B + G)/2. \quad (1)$$



Fig. 1 The schema of the LBP-method

3.2 *Local Binary Pattern*

The first approach is based on Local Binary Patterns texture descriptors. The LBP method, presented in [6], has a wide range of different applications. This method is based on the assumption that the texture has two complementary aspects, a pattern and its strength. In [7] a more generic, revised form of the LBP operator is presented, without limitations to the size of the neighborhood and to the number of sampling points. We have few different forms of LBP, such as: uniform LBP, rotation-invariant LBP or uniform rotation-invariant LBP available. The advantage of LBP is the efficiency of analyzing textures. The LBP method has a simple theory and combines properties of structural and statistical texture analyses methods. The paper [8] shows that it is favorable to use rotational invariant features for the analysis of the given biological structures, mostly anisotropic. The result of LBP for villi images is a noisy image. It is the result of an uneven distribution of stains. In order to obtain a coherent image, we have applied mean filter. Figure 1 presents scheme of LBP-method applied to trophoblast detection.

As a result, we have a coherent image which presents LBP features. The trophoblast is connected with the specified LBP features. That is why, we have applied the thresholding operation to the achieved LBP image. We have achieved an image which presents trophoblast surrounding the villi. Due to different thicknesses of the trophoblast, the image is not coherent. In order to obtain a coherent trophoblast image we have applied simple morphological operations, like erosion and dilation. Thanks to this, we have eliminated small elements and combined parts of the trophoblast. As a result, we have a coherent image presenting detected trophoblast surrounding the individual villi.

3.3 *Unser Features*

The second approach is based on modified formulas of Unser features [11]. It is connected with a normalized probability applied to the pixel intensity of an image. We have conducted research with the eight Unser features, such as: mean, variance, energy, correlation, contrast, homogeneity, cluster shade and cluster prominence. In defining texture descriptors, we have applied the histograms of the sum and difference of images [12]. These images are formed from the original image by applying the relative translation. An important step is to determine the image resolution and radius resulting in the best characterization of the local structures in specimens. Too small

Fig. 2 The schema of the Unser-method



radius causes little ambient impact of the properties of a given pixel and a strong heterogeneity. On the other hand, too large radius causes excessive generalization. We have applied the Unser features to detect the trophoblast, based on the schema in Fig. 2.

The first step is to calculate the Unser Features. Next, we have selected the best features which differentiate the trophoblast from the tissue. The trophoblast is connected with specified feature value. Therefore, we have applied thresholding operation to detect them. In order to achieve a coherent image of a trophoblast, we have used mathematical morphology. Thanks to this, we have removed small elements and combined parts of the trophoblast. As the result, we have the image of the detected trophoblast and the villi.

4 Results

The twelve cases of placental villi from spontaneous miscarriage were subject to the analysis. The developed methods were applied for trophoblast detection. Some examples of placenta villi and detected trophoblast are presented in the Fig. 3. The achieved results were compared with the expert results. Thanks to this, we could evaluate the algorithm's result. The quantitative analysis was based on the comparison areas of the trophoblast detected by the expert and by the algorithm.

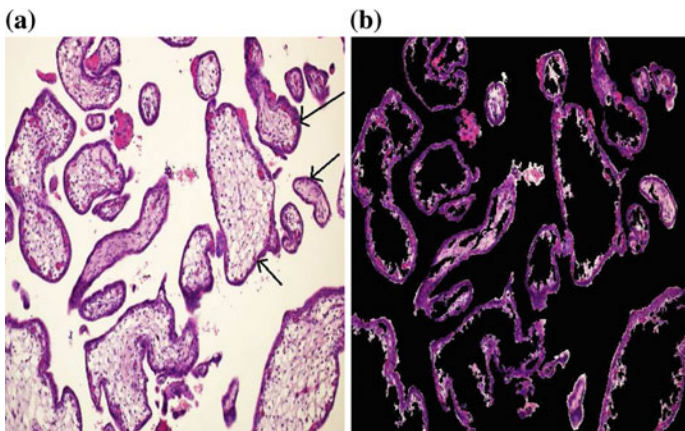


Fig. 3 The example of placenta villi (a), where arrows show trophoblast, and detected placenta trophoblast (b)

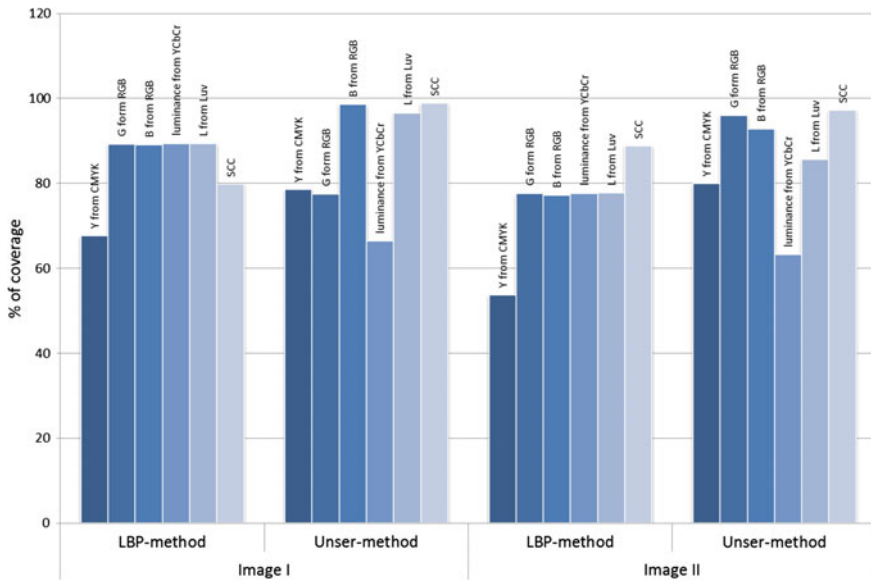


Fig. 4 The results of percentage of coverage for the trophoblast detected by the expert and the algorithm for five selected colour components and two images

In the first step, we have done the research of how the trophoblast surrounding the villi is distinguished in each of analyzed colour component. As a result, we have selected five colours components (Y from CMYK, G from RGB, B from RGB, Luminance from YCbCr, L from Luv, and SCC), where trophoblast is most distinguished. In Fig. 4, the percentage of coverage for the trophoblast detected by the expert and by the algorithm for five selected colour components and two images is presented.

Based on the research, we have selected two colour components, which have the highest percentage of coverage. We have selected colour components: L form Luv and SCC. In the next step of the trophoblast detection, the texture descriptors (LBP and Unser) were calculated for each specimen pixel. We have developed four presented algorithm versions:

- LBP-method for L colour component,
- LBP-method for SCC colour component,
- Unser-method for L colour component,
- Unser-method for SCC colour component.

The examples of selected features, after performing thresholding operation, and also the filtration operation in a case of LBP-method, are presented in Fig. 5. Differences between the algorithms are connected with the applied parameters. In the case of 'LBP-method for L colour component', the applied radius of the neighborhood region was 100 pixels. In the case of 'LBP-method for SCC colour component' we have used 5 as the radius of the neighborhood region. In the case of 'Unser-method

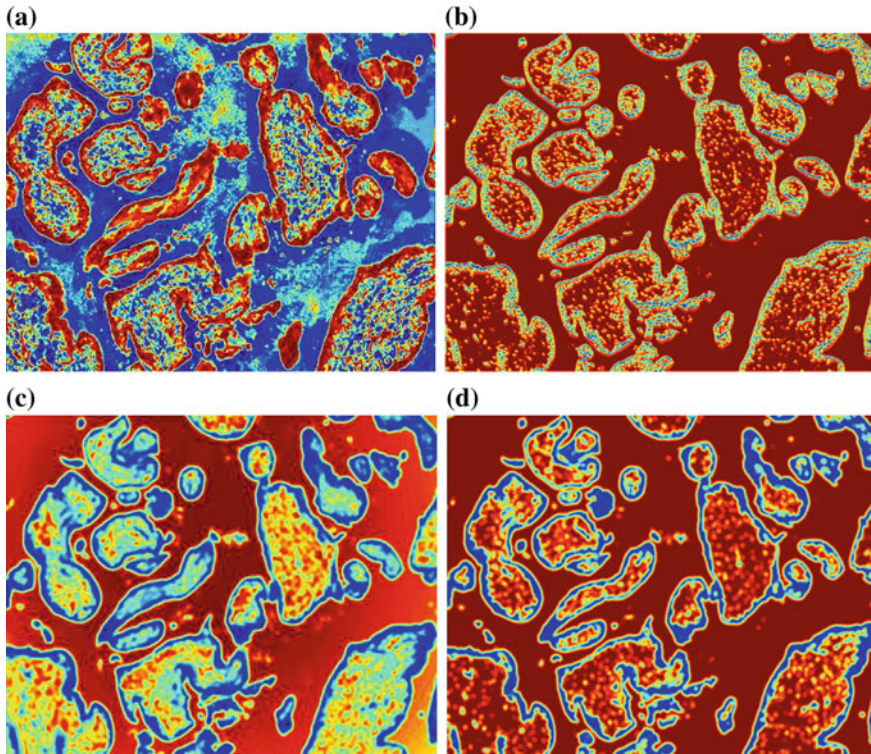
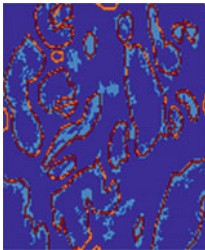
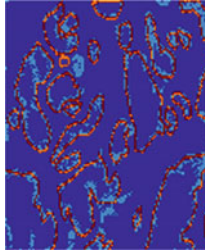
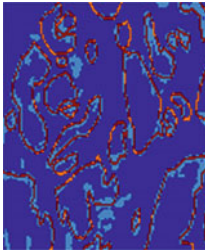
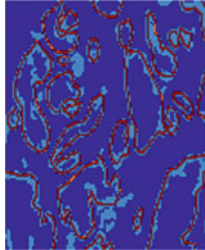


Fig. 5 The examples of selected features, where **a** LBP-method for L colour component, **b** LBP-method for SCC colour component, **c** Unser-method for L colour component, **d** Unser-method for SCC colour component

for L colour component’, we have selected the ‘mean’ feature (form Unser features), as the most distinguishing for the trophoblast. However, it is a regular mean function, which is also in a set of Unser features. In the last case (‘Unser-method for SCC colour component’) we have selected ‘homogeneity’ features, as the most distinguishing for the trophoblast. In all cases, we have applied simple mathematical morphology operations, as dilation and erosion [9]. The result of presented algorithms’ work is the detection of trophoblast surrounding villi. The examples of results are presented in Table 1.

The quantitative analysis was done to achieve villi maps. In order to assess the results, we have chosen parameter ‘% of coverage’, which presents percentage of trophoblast defined by the expert, which is included in the map detected by the algorithm. The results of specimen analysis are presented in Table 2 and Fig. 6. Based on them, we have observed that:

Table 1 An example of trophoblast detection for different methods

Method	Colour component	
	L from Luv	SCCI
LBP-method		
Unser-method		

- The best result has been achieved from the Unser-method for the SCC colour component. In this case, we have the highest percentage of covering trophoblast selected by the expert and trophoblast detected by the developed algorithm, while the difference between the maximum value and the minimum value is the smallest.
- The Unser-method results are characterized by higher mean values than the mean values for the LBP-method. However, in the case of L colour component, the difference between the maximum value and the minimum value is large and significant.
- In the case of the LBP-method, the best results were achieved for the L-colour component. The mean value is high and the difference between the maximum value and the minimum value is small.
- In the aspect of colour components, we can observe that SCC colour component gives good results for the Unser-method, while the L colour component gives good results for the LBP-method.

Our results confirm the effectiveness of automatic trophoblast detection. Figure 6 contains graph which presents statistical details (minimum, maximum, average) for the achieved results.

Table 2 The results of specimen analysis

Id/colour	LBP-method		Unser-method	
	L from Luv	SCCI	L from Luv	SCCI
1	80.84	74.62	94.44	97.85
2	89.31	79.77	96.56	98.83
3	81.05	74.51	91.39	95.30
4	77.70	74.44	85.63	97.28
5	84.90	77.86	94.61	99.83
6	74.68	66.34	67.84	95.60
7	69.28	68.47	76.42	94.96
8	64.93	69.33	80.25	94.69
9	80.67	53.51	57.31	78.50
10	76.71	76.26	87.65	97.14
11	75.68	63.71	62.10	87.17
12	69.00	62.01	78.53	86.70
Average	77.06	70.07	81.06	93.65
Median	77.20	71.89	82.94	95.45
Min	64.93	53.51	57.31	78.50
Max	89.31	79.77	96.56	99.83

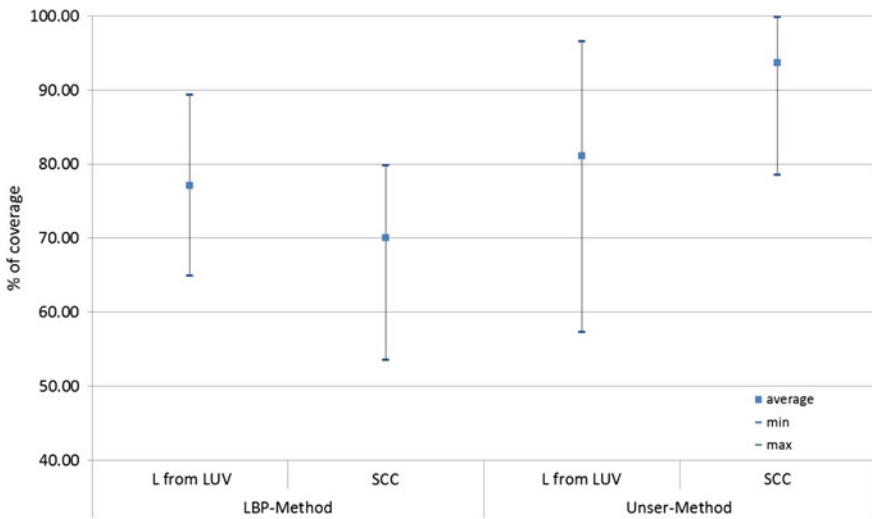


Fig. 6 The statistical details (min, max, average) for obtained results

5 Conclusions

We have applied two different texture descriptors for two different colour representations. Based on the achieved results, we have noticed that in all analyzed cases, the trophoblast surrounding villi was correctly detected. They have covered the trophoblast defined by the expert. It should be noted that the trophoblast surrounding villi can have different thicknesses, and that the collagen might surround them. The presented approach shows that based on LBP and Unser features for a selected colour component, we can detect trophoblast correctly. Our approach is characterized by simple application, short calculation time, and repeatability of results. The future research will include: collagen detection, hemorrhage detection, blood vessels detection, and development of parameters to evaluate placenta villi.

Acknowledgments This work has been supported by the National Science Centre (Poland) by the grant 2012/07/B/ST7/01203 in the years 2013–2016.

References

1. Almoussa, N., Dutra, B., Lampe, B., Getreuer, P., Wittmann, T., Salafia, C., Vese, L.: Automated vasculature extraction from placenta images. In: Proceedings of SPIE 7962 Medical Imaging (2011). doi:[10.1117/12.878343](https://doi.org/10.1117/12.878343)
2. Ayache, M., Khalil, M., Tranquart, F.: DWT to classify automatically the placental tissues development: neural network approach. *J. Comput. Sci.* **6**(6), 634–640 (2010)
3. Ernst, L.M., Gawron L., Fritsch M.K.: Pathologic examination of fetal and placental tissue obtained by dilation and evacuation. *Arch. Pathol. Lab. Med.* **137**, 326–337 (2013)
4. Hassan, T.M.M., Hegazy, A.M.S., Mosaed, M.M.: Anatomical and histopathologic analysis of placenta in dilation and evacuation specimens. *Forensic Med. Anatomy Res.* **02**(02), 17–27 (2014)
5. Maenpaa, T., Pietikainen, M.: Texture analysis with local binary patterns. In: Chen, C.H., Wang, P.S.P. (eds.) *Handbook of Pattern Recognition and Computer Vision*, pp. 197–216 (2005)
6. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recognit.* **29**(1), 51–59 (1996)
7. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
8. Sauer, C., Fehr, J., Ronneberger, O., Burkhardt, H., Saudau, K., Kurz, H.: Automated identification of large cell numbers in intact tissues—self-learning segmentation, classification, and quantification of cell nuclei in 3-d volume data via voxel-based gray scale invariants
9. Soille, P.: *Morphological Image Analysis. Principles and Applications*. Springer, Berlin (2003)
10. Stallmach, T., Hebisch, G.: Placental pathology: its impact on explaining prenatal and perinatal death. *Virchows Arch.* **445**(1), 9–16 (2004)
11. Swiderska, Z., Markiewicz, T., Grala, B., Kozłowski, W.: Texture and mathematical morphology for hot-spot detection in whole slice images of meningiomas and oligodendrogliomas. In: 16th International Conference, CAIP 2015, Valletta, Malta, 2–4 Sept LNCS 9257, pp. 1–12 (2015)
12. Unser, M.: Sum and difference histograms for texture classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 118–125 (1986)

Machine Learning Can Reliably Distinguish Histological Patterns of Micropapillary and Solid Lung Adenocarcinomas

Nathan Ing, Sadri Salman, Zhaoxuan Ma, Ann Walts,
Beatrice Knudsen and Arkadiusz Gertych

Abstract The World Health Organization recommends subclassification of lung cancer according to the percentages of histologic subtypes within a tumor. The manual quantification of lung tumor composition is very time consuming, but it can potentially be aided by a machine learning application. We have updated our previously developed methodology to segment and distinguish solid and micropapillary lung tumor subtypes. Binary tumor masks delineated by machine learning were defined by the mean area of binary objects and by the number of objects found in an image frame. These two features distinguished solid ($n = 31$) and micropapillary ($n = 61$) histologic subtypes with excellent performance ($p < 4.04E-19$) for three different frame sizes. Our method to quantify tumor growth patterns applied to histological images of lung adenocarcinoma, demonstrates for the first time that it is feasible to quantify the composition of histological subtypes in individual lung cancers.

Keywords Machine learning · Image analysis · Lung cancer · Tissue classification · Tumor quantification

1 Introduction

Lung cancer is currently the leading cause of cancer-related deaths worldwide. In the U.S., the American Cancer Society estimates there will be 221,200 new cases of non-small cell lung cancer (NSCLC) and 158,400 associated deaths in 2015 [19].

N. Ing · Z. Ma · B. Knudsen
Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles,
CA 90048, USA

A. Walts · B. Knudsen · A. Gertych
Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center,
Los Angeles, CA 90048, USA

N. Ing · S. Salman · A. Gertych (✉)
Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
e-mail: arkadiusz.gertych@cshs.org

Survival varies with histological type of carcinoma and tumor stage at the time of diagnosis. Despite advances in detection and therapy, the 5-year overall survival for NSCLC remains low. Adenocarcinoma is currently the most common type of lung cancer and accounts for about 50% of NSCLC [19]. These tumors occur in smokers and former smokers, and are the most common type of lung cancer in never-smokers. Adenocarcinomas of the lung are more common in women than in men and occur more frequently in younger persons than other types of lung cancer.

The new 4th edition (2015) of the World Health Organization (WHO) classification of tumors of the lung and other thoracic organs recognizes five histological patterns of invasive adenocarcinoma of the lung: lepidic, papillary, micropapillary, acinar, and solid and recommends that tumors be subclassified based on the predominant growth pattern exhibited in the resected tumor [15]. Multiple recent studies showing differences in prognosis, benefit of adjuvant therapy, risk of recurrence after wedge or segmental resection, and/or frequency of targetable mutations among these histological subtypes support this recommendation [7, 11, 16, 17, 20, 22]. It has also been suggested that the cribriform pattern be added to this classification [4]. Assessment of the histological patterns in a tumor by a pathologist requires a composite evaluation of subjective “eyeball” estimations of the percentage of each pattern in each of several slides prepared from the tumor. The presence of mixed histological patterns in more than 80% of lung adenocarcinomas [13] makes it particularly difficult for pathologists to comply with the International Association for the Study of Lung Cancer, American Thoracic Society, and European Respiratory Society (IASLC/ATS/ERS) recommendation that the presence of each histological pattern be reported in 5% increments [14], a recommendation supported by studies that suggest the prognostic value of this classification is not limited only to the predominant growth patterns [4, 6]. Thus, there is a potential role for development of image analysis algorithms that can accurately identify and reproducibly quantitate these histological patterns. Computerized quantitation of tumor patterns would decrease intra- and inter-observer variability in tumor sub-classification. Accurate and reproducible stratification of patients is a prerequisite for further research in this area and optimal individualized clinical management.

This study was designed to explore the potential of computer image analysis to identify, distinguish, and quantitate histological patterns of pulmonary adenocarcinoma. As micropapillary and solid patterns have consistently been shown to be associated with poorer prognosis [16, 17, 22], we chose these two patterns for this pilot study.

Since no image analysis methods have been developed for this purpose so far, and it is unclear how these patterns could be quantitated, we propose the following analytical pipeline. First, we apply an image segmentation method that separates cancer cells from other components of the tumor microenvironment (inflammatory cells and stroma) to arrive at a binary tumor mask. Second, we characterize the tumor mask using two features: the mean area of objects and the number of binary objects detected in an image frame. These features are known as binary texture descriptors and were previously applied to quantitate conformation of chromatin in the cell

nucleus [2, 9, 21]. Third, we demonstrate the performance of both steps and show that binary masks of solid and micropapillary tumors can be reliably distinguished by these two features.

2 Materials and Methods

2.1 Data Acquisition

Four hematoxylin (H) and eosin (E) stained slides of lung adenocarcinoma from two cases of predominantly solid and two cases with predominantly micropapillary patterns were obtained from the pathology archives at our institution. Slides were reviewed by a pathologist who manually circled tumor areas and labeled each tumor pattern. Slides were digitized at $20\times$ magnification ($0.5\ \mu\text{m}/\text{pixel}$) using a whole slide scanner Aperio Turbo AT (Leica Biosystems) equipped with a clinical grade RGB camera. The manually delineated regions of interest were extracted from the digitized slides. A total of 91 H&E tumor images (60 micropapillary and 31 solid) were collected and exported as 1200×1200 pixel matrices in the 24-bit tiff format. These original H&E images were subsequently divided using regular 2×2 and 3×3 grids. The original images and subimages of 600×600 ($n = 240$ for micropapillary and $n = 124$ for solid tumors) and 400×400 pixels ($n = 540$ for micropapillary and $n = 279$ for solid tumors) were obtained through this division and used to extract features for tumor mask characterization.

2.2 Methods Overview

Our main goal was to develop a method that is capable of delineating large and compact areas of solid tumor as well as small, isolated papillary fronds that are characteristic of the micropapillary pattern (Fig. 1). We therefore revised and upgraded our previously published tissue classification technique [3, 12] to properly capture the geometry of the two tumor patterns with a single method.

This methodology involves intensity histograms (IH) that relate to the specific affinity of hematoxylin to DNA. Cancerous nuclei often exhibit open chromatin structures which appear diffusely stained by hematoxylin with darkly stained nucleoli. This feature is reflected in IH histograms and separates the cancerous epithelium from the stroma. However, inflammatory cells that are present in the stroma or epithelium may occur in large quantities and therefore should be separated from these two tissues by image analysis approaches. These cells are smaller than a typical tumor cell and their nuclei stain densely with hematoxylin.

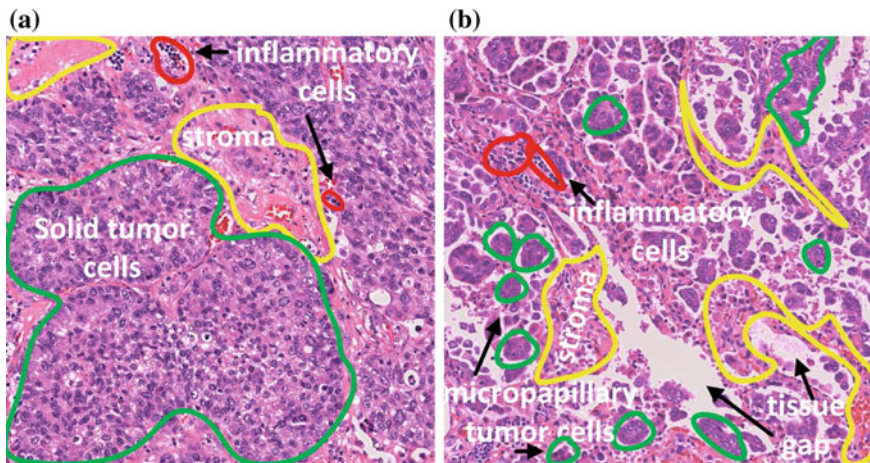


Fig. 1 Example histological patterns of lung adenocarcinomas: **a** solid, **b** micropapillary. Example regions of histologic structures important to this analysis are labeled

2.3 Preprocessing

Hematoxylin binds and stains DNA in purple-blue, and eosin binds to proteins in cytoplasm and stains them in pink. Since these stains adhere to molecular structures and obey the Beer-Lambert law it is possible to isolate and quantify the amount of each stain. To correct for H&E staining variations and color distribution, the images were first color normalized in the LAB color space [8]. We then applied the color deconvolution procedure [10], to digitally separate images of hematoxylin and eosin from the color normalized H&E image. Similarly to our previous studies, we utilize the differential information embedded in the intensity levels of eosin and hematoxylin to classify areas of the lung tissue and arrive at discrete masks of the cancer, stroma and inflammatory cells.

2.4 Segmentation

Our analytical framework takes advantage of methodologies developed for the quantification of prostate cancer tissue [3, 12] with some important modifications. Briefly, a Random Forest classifier (RF) classifies pixels into either two or three classes based on that pixel's local intensity histograms derived from color-deconvoluted hematoxylin and eosin images. A three-dimensional (3-D) intensity histogram (IH) image is formed first. Its x and y dimensions are the same as in the input image. Values along its z dimension are filled by IHs found by window W sliding in parallel through H and E images so that one 18-bin hematoxylin and one 18-bin eosin

histogram are obtained. The window W is stepped across and down the image at regular intervals, and at each step storing the combined 36-bin local histograms of H and E intensities called a descriptor. At the end of the window stepping procedure, bicubic interpolation fills the 3-D IH image at intermittent image coordinates which did not happen to fall at the center of any W . Essentially, a 36-bin descriptor corresponds to a pixel (x, y) in the original H&E image. If multiple sizes of W are implemented, the descriptors will have the capacity to characterize tissues at different image resolutions [3]. Thus, for the purpose of this study we set the size of W to be $W = 64$ and $W = 16$ in order to properly capture tumor cells in heterogeneous tissue areas with high or low cell densities. Respective stepping intervals are $W/4$ and $W/8$.

A two-step classification was developed utilizing the descriptors determined by the 64-pixel window (STEP-1), and 16-pixel window (STEP-2). At each STEP a RF classifier with 35 decision trees was applied to classify respective descriptors. The RF classifiers were trained with descriptors extracted from windows W that were manually selected by the pathologist. Specifically, 30 descriptors were selected for each class to train the RF classifier separating cancer and stroma in STEP-1, and 30 each for the RF classifier separating cancer cells, from inflammatory cells and stromal cells in STEP-2. 150 such descriptors were collected and then box-plotted (Fig. 2) to illustrate their properties. In our previous publications [3, 12] we demonstrated that differences in shape and distribution of the intensity-based descriptors lead to a robust separation of stromal and epithelial areas in prostate cancer tissue. Based on the previous results, we reasoned that the descriptors extracted from lung cancer tissue and collected for the classification STEP-1 (Fig. 3a) and STEP-2 (Fig. 3b) will possess similar discriminating power. Thus, the RF classifier in STEP-1 is expected to segment cancer from stromal regions, whereas that in STEP-2 segments cancer cells from inflammatory cells and other stromal components. Note that the two-class classification involving the 64-pixel window is better tailored for segmenting large and homogenous tumor areas, whereas three-class classification with the 16-pixel window is suitable for segmenting individual cells in a heterogeneous cluster (Fig. 3).

Classification results are then mapped back to their original image coordinates, forming a coherent segmentation mask of the structures in question for each of the STEPs. Our scheme was applied in a dual-scale sense (64 and 16 pixel windows) and results overlaid in an order determined *a posteriori* (Fig. 3). A combination of these masks gives an image whose pixels are classified as one of the three histologic categories. To obtain a final result, the intermediate masks from STEP-1 and STEP-2 classifications are sequentially merged. A logical OR operation was applied to combine stromal masks in the first step of merging. In the two remaining merges the cancer and inflammatory cell masks were respectively overlaid (pixel labels were replaced) on the mask obtained from the preceding step. The segmentation performance was assessed by two raters who assigned each image a score from 1 (unacceptable) to 5 (excellent) using semi-quantitative scores (Table 1).

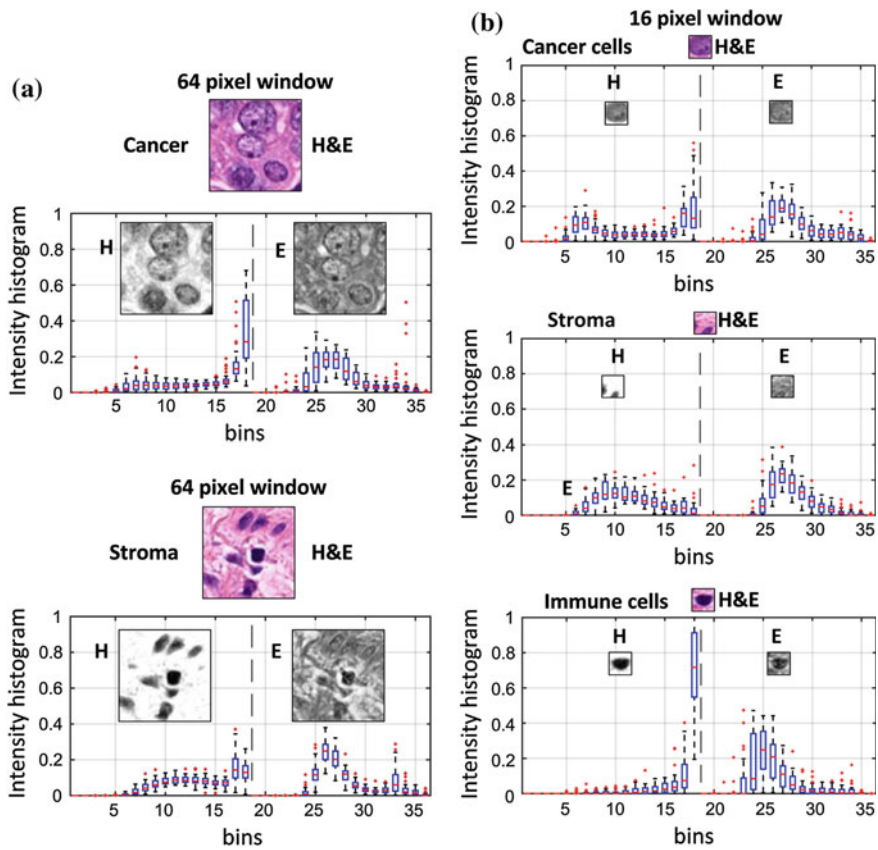


Fig. 2 Box plots of hematoxylin (H) and eosin (E) intensity histograms. The intensity histograms were extracted by means of **a** 64 pixel, and **b** 16 pixel scanning windows. Respective micrographs show example cancer and stromal areas as well as individual cells selected for training of classifiers to detect cancer, stroma and inflammatory cells. Box plots demonstrate how the 64 and 16 pixel windows capture local tissue characteristics

2.5 Distinguishing Subtypes of Adenocarcinoma by Discrete Objects and Area

Segmented tumor was represented as binary masks: 1–tumor, 0–all other classes (Fig. 4, bottom row). The binary masks were post-processed. A binary opening with a disk shaped structuring element with a radius of 5 was applied to clean up spurious shapes in the masks and smooth mask edges for subsequent quantification. We then chose two basic chromatin conformation features described in [2, 9, 21] to characterize the morphological difference between the micropapillary and solid tumor patterns. Our feature selection was guided by qualitative standards recently established by WHO [15, 19]. The output binary masks were quantitated by two features: the number of objects (NO) with 8-connectivity in a given frame, normalized by the

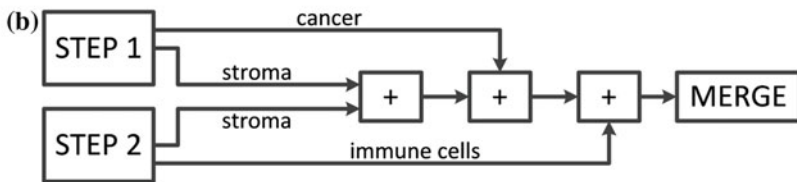
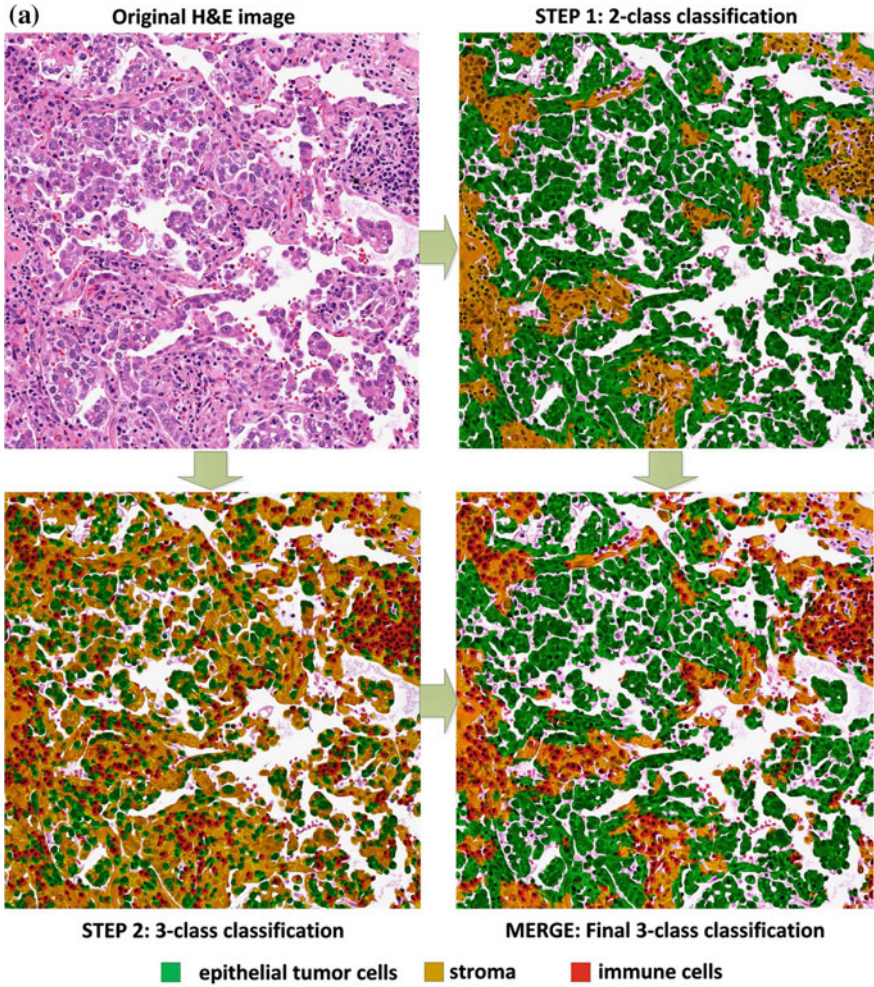


Fig. 3 Image classification workflow. Pixels of color-deconvoluted H&E images are classified by a random forest classifier in two steps. **a** in STEP 1 cancer (green) and non-cancer (yellow) areas pertain to the classification with the 64 pixel window. In STEP 2 the classification is carried out using a 16 pixel window and detects cancer cell nuclei (green), inflammatory cell nuclei (red) and stromal components (yellow). Cancer, stromal and inflammatory cell masks are combined in a predefined order **b** to obtain a delineation of the whole tumor. The main goal of this approach is to preserve the histological pattern of tumor for subsequent analyses

Table 1 Scores and criteria of segmentation quality

Score	Description
1 (Unacceptable)	Seemingly random distribution of classes about the image. More errors than correct classifications
2 (Poor)	Multiple gross misclassifications e.g. inflammatory cells classified as tumor, especially missing large areas of tumor
3 (Average)	Some small false positive tumor areas, vast majority of tumor structures are detected but not necessarily perfectly segmented
4 (Good)	No false positive tumor areas. Immune cell groups and individual cells are well detected. Close but separate tumor structures may be connected by thin “bridges”. All tumor structures detected and well segmented
5 (Excellent)	Near perfect detection and segmentation. It would be challenging to manually generate more accurate delineations

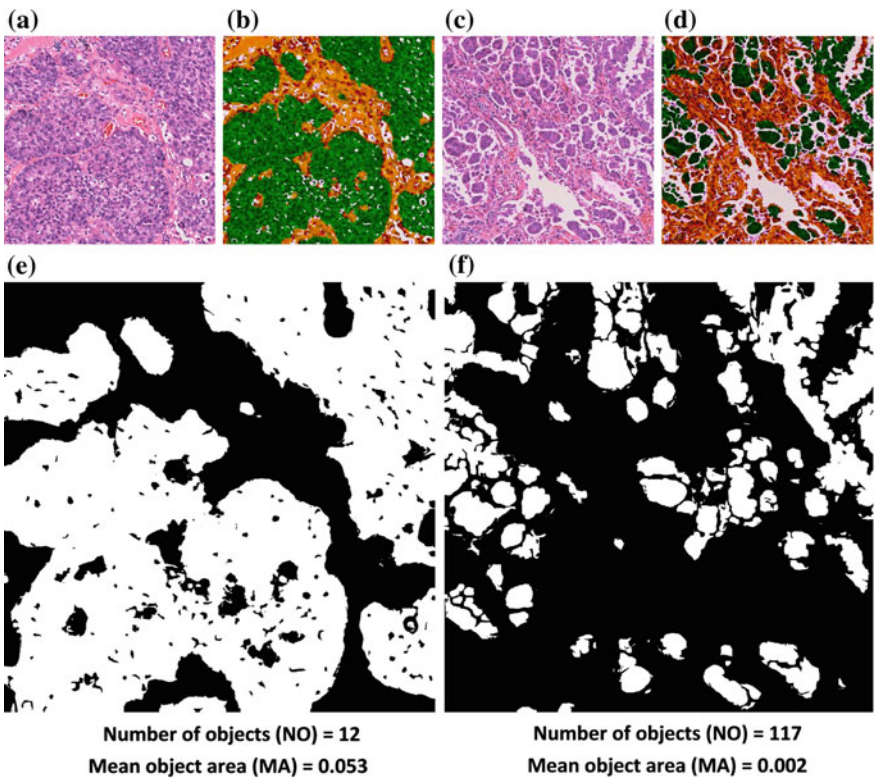


Fig. 4 Processing results of solid (a) and micropapillary (c) tumors from Fig. 1. Respective 3-class image classification results (b and d) obtained by our workflow in Fig. 3 yielded an accurate delineation of tumors (green color). Solid tumor masks analyzed by our approach have an overall smaller number of binary objects (NO) (e) than the masks of micropapillary tumors (f) which on the other hand have smaller mean area of objects (MA)

area of the frame and the mean area of these objects (MA), also normalized by frame area. To demonstrate the effect of image frame size on the MA-NO features, the original images were divided. Besides the original image (1200×1200 pixels), subsequent images were generated as by dividing the original into four parts (600×600 pixel image frame), and into nine parts (400×400 pixel image frame). MA and NO were calculated for each frame at all sizes and statistically evaluated by the t-test.

3 Results

3.1 Micropapillary and Solid Pattern Distinction

Normalized mean MA and NO with standard deviations from the analysis for different image size and statistical evaluations are outlined in Table 2. Co-distributions of normalized MA and NO features are plotted on log-log scale (Fig. 5). For the

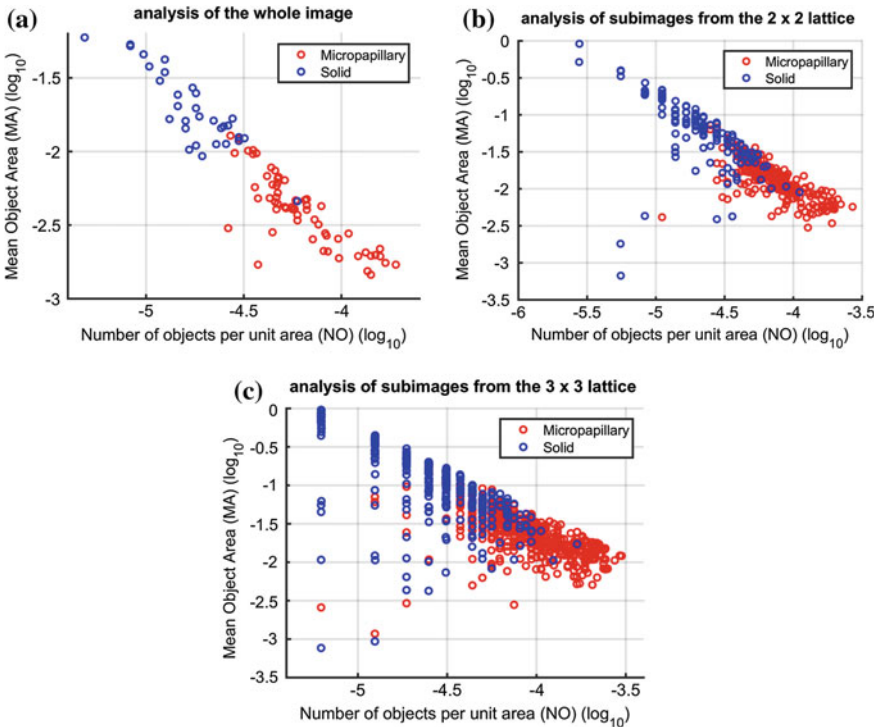


Fig. 5 Co-distributions of mean object area (MA) and number of objects (NO) in binary tumor masks (Fig. 4). MA and NO were derived from **a** whole images (1200×1200 pixels), **b** medium size sub-images (600×600 pixels), and **c** small sub-images (400×400 pixels). Raw values of MA and NO were normalized respectively by the size of the whole image or a sub-image. The separation of solid and micropapillary tumor masks was excellent for all image sizes

Table 2 Mean number of objects (NO) and mean area of objects (MA) features in images containing solid and micropapillary tumor patterns

	Mean number of objects (NO)		
	1200 × 1200	600 × 600	300 × 300
Micropapillary	102 ± 58	29.9 ± 16.8	15.5 ± 8.7
Solid	28.1 ± 14.8	9.7 ± 6.3	5.4 ± 3.8
<i>p</i> -value ($\alpha = 0.5$)	4.04E-19	2.81E-53	1.20E-95
	Mean area of objects (MA)		
	1200 × 1200	600 × 600	300 × 300
Micropapillary	6.7E2 ± 4.0E2	6.0E3 ± 4.2E2	5.3E2 ± 4.6E2
Solid	3.2E4 ± 2.1E4	3.2E4 ± 4.0E4	2.8E4 ± 3.3E4
<i>p</i> -value ($\alpha = 0.5$)	6.95E-21	5.86E-40	2.73E-70

NO and MA were compared in three image sizes. Means are reported from raw, un-normalized quantities

Table 3 Tumor segmentation scores. Numbers represent the mean score ± standard deviation per rater

	Rater 1	Rater 2
Micropapillary ($N = 60$)	3.37 ± 1.02	3.81 ± 0.70
Solid ($N = 31$)	3.48 ± 0.84	3.90 ± 1.04
Micropapillary and solid ($N = 91$)	3.4 ± 0.91	3.84 ± 0.82

400 × 400 and 600 × 600 sizes, some sub images contained no detected tumor area. These were excluded from analysis (Table 2).

3.2 Segmentation Performance

Performance was assessed according to Table 1 independently by two raters. The mean score for each subtype and rater is shown in Table 3.

4 Discussion

Histological subtypes of lung adenocarcinoma have recently gained attention as a possible prognostic indicator. The composition of subtypes in a cancer is currently estimated in 5% increments across several slides of resected tumor. However, this process is wrought with inter and intra-observer inconsistency. Computer vision methods have the potential to remove observer-related variability by precisely

classifying each subtype and by accurately measuring its percentage in the total tumor. A key advantage of this approach is reproducibility at frame sizes not feasible for manual annotation or analysis. The methods we developed for subtyping of lung adenocarcinoma overcome confounding factors that are often encountered in histologic image data.

4.1 Tumor Microenvironment Segmentation

Tumors display varying growth patterns of cancerous cells that are embedded in stroma. The stroma is composed of multiple cell types, which include inflammatory cells infiltrating and at times encasing the tumor. The segmentation is performed with small windows—a way that is blind to the relatively macroscopic differences in overall tumor morphology and arrangement. In contrast, differentiation of tumor pattern may only be possible using whole images or subimages that can enclose large fragments of tumor morphology. We devised a method that can automatically find cancer and inflammatory cells and potentially other cell types and structures in H&E images. It classifies IH histograms obtained with various windows, and combines the results based on the strength of segmentation at each frame size which is a step towards comprehensive segmentation of the tumor microenvironment in H&E images. We found that with the window size of $W = 64$, tumor and stromal areas were well separated by IHs. For this window size, however, inflammatory cells have local IHs similar to tumor cells and were often misclassified. Additionally, tumor clusters are occasionally separated by narrow strips of stroma, which were often wrongly classified, especially if the stromal width was less than the scanning window size. We therefore introduced a smaller window size, $W = 16$, specifically to accomplish better isolation of inflammatory cells and to arrive at a more accurate tumor mask.

Another parameter that impacts the classification is the spacing between the centers of consecutive windows, which we call the stepping interval, defined as $(W/4$ and $W/8)$ respectively for $W = 64$ and $W = 16$ of window sizes. The stepping is the distance, in pixels, that the window slides between IH calculations. IHs at pixels in between are interpolated. If the stepping interval is a small fraction of the scanning window, then fewer IHs need to be interpolated. Since changes in cellular composition across the tissues cause variation in local staining intensities, the interpolation is not guaranteed to produce an accurate classification when structures (tumor, inflammation, stroma) collide. A smaller scanning window yields a more locally contained IH, excellent for detecting small objects. A stepping interval $W/8$ (2 pixels) was selected for $W = 16$, to work with a three-class RF. This made the capture of inflammatory cells and intermittent stroma accurate. Although we did not systematically evaluate the influence of the stepping interval on the segmentation results, we took advantage of our previously published data to guide its selection. In [3] we showed that a classification of non-interpolated IHs (stepping interval = 1 pixel) is generally more accurate than that involving interpolated IHs with stepping interval = 16 pixels.

However, the accuracy gained from a small stepping interval was disproportionate to the increased calculation time. Since the segmentation results for a two-pixel and one-pixel stepping interval for $W = 16$ pixels were nearly identical, we selected the former to shorten the overall analysis time.

Since the inter-slide variability in the staining intensity of hematoxylin and eosin has the potential to decrease the segmentation performance, we elected to train the algorithm individually on each case that we analyzed. In this set-up the same algorithm performed comparably well in both tumor patterns. The segmentation performance was consistent across specimens, and, collectively, the combined classification improved the tumor area segmentation over a single-window method. Segmentation quality in the majority of images was predominantly average or good. The raters found respectively 13 and 3 images to be segmented as poor or unacceptable and 8 to be excellent. The mean scores of Rater 1 and 2 were 3.4 and 3.84. Since a significant manual effort is required to precisely delineate the tumor we only performed a semi-quantitative assessment of the segmentation. The high scores collectively given by two independent raters encourage us to continue the development of this approach and undertake more objective and rigorous evaluations with immunohistochemically generated ground truth masks of the tumor microenvironment.

4.2 Tumor Patterns Are Separable with Similar Accuracy at Different Analytical Dimensions

The MA and NO features used to describe micropapillary and solid tumor patterns were inspired by how they are recognized by a pathologist. Solid tumor masses, when prepared on a slide, appear as sheets of cancer cells, without intervening gaps in tissue integrity or stromal tissue. Micropapillary tumors, on the other hand, appear as disjointed, small clusters of cancer cells, separated by clear spaces and arranged around a stromal stalk. MA and NO were distinct in binary masks of these two patterns (p -value < 4.04E-19), and the classification performance was not affected by the image size. In contrast to other features characterizing the binary texture of images, which are often abstract, both MA and NO are directly related to features that encompass the qualitative description of the tumor pattern provided by WHO [15, 19].

In addition to the methods that we specifically developed for the identification of subtypes of lung adenocarcinoma, multiple other methods exist that can be used to quantify tumor growth pattern from spatial properties of the binary tumor mask. For example the box-counting fractal dimension and sliding-box lacunarity have been previously applied to distinguish tumor growth patterns in binary nuclear masks or intensity images of prostate cancers [1, 5, 18, 21]. They enumerated space-filling

and gappiness of objects against a background. However, the fractal dimension and lacunarity have not yet been evaluated on tumor masks which include individual and clustered tumor cells.

Our framework may be feasibly reapplied to quantify any tumor growth pattern within a complex microenvironment. We have not yet explored the limits of permissible frame size. Preliminary data suggests the MA-NO features are robust for the three frame sizes that we chose and plan to resolve this issue in future studies.

5 Conclusions

We extended the methods that were previously developed for analysis of prostate cancer to include more elements from the tumor microenvironment. Here we apply the updated method to the histologic subtyping of lung adenocarcinoma and demonstrate that it properly captures the shapes of solid and micropapillary tumors, is robust, and is computationally inexpensive. The mean area (MA) and number of connected objects (NO) constitute a novel and attractive set of features to separate and quantify the percentages of histological patterns within lung adenocarcinomas that are of prognostic relevance. In addition to further testing the clinical significance of the new features in the subtyping of lung adenocarcinomas in a large cohort, we will seek opportunities to evaluate the performance of our method on masks derived from lepidic, papillary and acinar tumor growth patterns.

References

1. Baish, J.W., Jain, R.K.: Fractals and cancer. *Cancer Res.* 15, **60**(14) 3683–3688 (2000)
2. Gann, P.H., Deaton, R., Amatya, A., Mohnani, M., Reueter, E.E., Yang, Y., et al.: Development of a nuclear morphometric signature for prostate cancer risk in negative biopsies. *PLoS ONE* **8**(7), e69457 (2015)
3. Gertych, A., Ing, N., Ma, Z., Fuchs, T.J., Salman, S., Mohanty, S., Bhele, S., Velásquez-Vacca, A., Amin, M.B., Knudsen, B.S.: Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Comput. Med. Imaging Graph.* **46**, 197–208 (2015)
4. Kadota, K., Yeh, Y.C., Sima, C.S., et al.: The cribriform pattern identifies a subset of acinar predominant tumors with poor prognosis in patients with stage I lung adenocarcinoma: a conceptual proposal to classify cribriform predominant tumors as a distinct histologic subtype. *Mod. Pathol.* **27**, 690–700 (2014)
5. Ma, Z., Xiaopu, Y., Amin, M., Knudsen, B., Gertych, A.: Fractal descriptors accurately distinguish between growth patterns of prostate cancers. *Lab. Invest.* **95**, 399–399 (2015)
6. Mäkinen, J.M., Laitakari, K., Johnson, S., et al.: Nonpredominant lepidic pattern correlates with better outcome in invasive lung adenocarcinoma. *Lung Cancer* **90**(3), 558–74 (2015)
7. Nitadori, J., Bograd, A.J., Kadota, K., et al.: Impact of micropapillary histologic subtype in selecting limited resection vs lobectomy for lung adenocarcinoma of 2 cm or smaller. *J. Natl. Cancer Instit.* **105**, 1212–1220 (2013)
8. Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Graphics Appl.* **21**(5), 34–41 (2001)

9. Rodenacker, K., Bengtsson, E.: A feature set for cytometry on digitized microscopic images. *Anal. Cell. Pathol.* **25**(1), 1–36 (2003)
10. Rulfrok, A.C., Johnston, D.A.: Analytical and quantitative cytology and histology. *Int. Acad. Cytol. Am. Soc. Cytol.* **23**(4), 291–299 (2001)
11. Russell, P.A., Barnett, S.A., Walkiewicz, M., et al.: Correlation of mutation status and survival with predominant histologic subtype according to the new IASLC/ATS/ERS lung adenocarcinoma classification in stage III (N2) patients. *J. Thorac. Oncol.* **8**, 461–468 (2013)
12. Salman, S., Ma, Z., Mohanty, S., Bhele, S., Chu, Y-T., Knudsen, B., Gertych, A.: A machine learning approach to identify prostate cancer areas in complex histological images. In: Pietka, E., Kawa, J., Wieclawek, W. (eds.) *Information Technologies in Biomedicine*, vol. 3, pp. 295–306 (2014)
13. Travis, W.D., Brambilla, E., Müller-Hermelink, H.K. (eds.) *Pathology and Genetics of Tumors of the Lung, Pleura, Thymus and Heart. WHO Classification of Tumors*, vol. 10, 3rd edn. WHO Press, Geneva Switzerland (2004)
14. Travis, W.D., Brambilla, E., Noguchi, M., et al.: International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. *J. Thorac. Oncol.* **6**, 244–285 (2011)
15. Travis, W.D., Brambilla, E., Burke, A.P., Marx, A., Nicholson, A.G. (eds.): *WHO Classification of Tumors of the Lung, Pleura, WHO Classification of Tumors*, vol. 7, 4th edn. WHO Press, Geneva Switzerland (2015)
16. Tsao, M.-S., Marguet, S., Le Teuff, G., et al.: Subtype classification of lung adenocarcinoma predicts benefit from adjuvant chemotherapy in patients undergoing complete resection. *J. Clin. Oncol.* **33**, 3439–3446 (2015)
17. Tsuta, K., Kawago, M., Inoue, E., et al.: The utility of the proposed IASLC/ATS/ERS lung adenocarcinoma subtypes for disease prognosis and correlation of driver gene alterations. *Lung Cancer* **81**, 371–376 (2013)
18. Waliszewski, P., Wagenlehner, F., Gattenlöhner, S., Weidner, W.: On the relationship between tumor structure and complexity of the spatial distribution of cancer cell nuclei: a fractal geometrical model of prostate carcinoma. *Prostate* **75**(4), 399–414 (2015)
19. www.cancer.org/cancer/lungcancer. Accessed 3 April 15
20. Yoshizawa, A., Motoi, N., Riley, G.J., et al.: Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. *Mod. Pathol.* **6**, 1496–1504 (2011)
21. Young, I.T., Verbeek, P.W., Mayall, B.H.: Characterization of chromatin distribution in cell nuclei. *Cytometry* **5**, 467–474 (1986)
22. Zhang, Y., Li, J., Wang, R., et al.: The prognostic and predictive value of solid subtype in invasive lung adenocarcinoma. *Sci. Rep.* (2014). doi:[10.1038/srep07163](https://doi.org/10.1038/srep07163)

Apoptosis in Cancer Cells

Eva Blahovcová, Henrieta Škovierová, Ján Strnádel,
Dušan Mištuna and Erika Halašová

Abstract Programmed cell death changes and loss of cell growth and anti-growth regulation signals may result in uncontrolled cell proliferation, disorganized tissue growth and tumor formation. Malignant solid tumors, i.e. lung and brain tumors, are one of the most common and the most serious diseases nowadays and therefore considered a major issue in modern medicine. Despite intensive progress in cancer research, the survival rate of patients with diagnosed malignant solid tumors hasn't improved significantly. Radiotherapy and chemotherapy work predominantly by inducing apoptosis. The purpose of this article is to describe the role of apoptosis and to review currently known agents that allow induction of apoptosis in cancer cells. The identification of potential new targets for tumor treatment and processes in altered tumor cell metabolism is the challenging area of research. Recognizing the attributes of malignant tumor growth inhibition may help in the development of novel treatment strategies.

Keywords Apoptosis · Cancer · Solid tumor · Apoptotic inhibitors

1 Introduction

Cancer is very heterogeneous and represent one of the most common and the most serious group of diseases nowadays. Specific aberrant expression of multiple genes, disturbances in tissue homeostasis and deregulation of growth cycle's signal cascade

E. Blahovcová (✉) · H. Škovierová · J. Strnádel · E. Halašová
Jessenius Faculty of Medicine in Martin, Department of Molecular Medicine,
Biomedical Center Martin, Comenius University in Bratislava, Martin, Slovakia
e-mail: evablahovcova@yahoo.com

D. Mištuna
Jessenius Faculty of Medicine in Martin, Clinic of Surgery and Transplant Center,
Comenius University in Bratislava, Martin, Slovakia

E. Halašová
Jessenius Faculty of Medicine in Martin, Department of Medical Biology,
Comenius University in Bratislava, Martin, Slovakia

may result in tumor formation and progression. The treatment of solid tumors depends on the histological diagnosis. The mainstay of therapy is the resection of tumor followed by postoperative radiotherapy and/or chemotherapy, depending on type, stage, malignancy grade and individual patient's acute or late responses to treatment. The ability of the cell to avoid programmed cell death is one of the major features in tumor formation [13]. The deregulation of apoptotic mechanism plays an important role in the progression and pathogenesis of cancer. In fact, radiotherapy and chemotherapy should induce apoptosis. But the resistance to conventional anti-cancer therapy is a major obstacle to successful treatment [35]. Selective inhibitors of apoptotic signaling pathway provide option for studying of apoptotic proteins and their role in resistance to apoptotic mechanism in tumor cells. Moreover apoptotic inhibitors provide new possibilities in solid tumors treatment, including lung or brain cancer. Despite a progress in cancer treatment and diagnostic methods, the prognosis for patients with diagnosed malignant solid tumor is still extremely serious and makes tumor disease a challenging area of research [34].

2 Apoptosis

Evasion of programmed cell death is a key hallmark in tumor formation [13]. Apoptosis is a common property of multicellular organism, plays a fundamental role in many physiological processes and is present in all cell types throughout the body [32]. It is also one of the major mechanisms in cell homeostasis preservation in tissues, and immune system [26]. Apoptosis can be activated by multiple pathways that differ among tissue type and pathological conditions. The extrinsic pathway is triggered by activation of cell surface receptors, while intrinsic pathway is activated by a variety of extracellular and intracellular stresses and associated with mitochondria perturbation [8]. Convergent signals in proteolytic cascade of both apoptotic pathways activate caspases, family of cysteinyl-proteases, which play the key role in the initiation and execution of programmed cell death [28].

Extrinsic apoptotic pathway is activated by the stimulation of the tumor necrosis factor receptor superfamily located on the cell surface. After ligand binding, death domain (DD) of receptors attracts the intracellular adaptor protein, which recruit procaspase 8 and form death-inducing signaling complex (DISC). In relationship to these events, caspase 8 and subsequently the major effector caspase 3 are activated [22, 24].

Intrinsic apoptotic pathway is activated after different conditions of cellular stress, including non-reparable DNA damages, oxidative stress, irradiation, hypoxia, protein p53, high cytoplasmic Ca²⁺ concentration and treatment with cytotoxic drugs. The mitochondrial dependent pathway is mediated by Bcl-2 (B-cell lymphoma-2) apoptotic proteins and by inhibitors of apoptosis proteins (IAPs) family [8, 22].

Mitochondria plays a crucial role in the initiation of intrinsic apoptotic pathway. Death-promoting stimuli favors mitochondrial membrane permeabilization which leads to releasing of intermembrane space proteins (such as cytochrome c, second

mitochondria-derived activator of caspase (Smac), apoptosis inducing factor (AIF), endonuclease G) into the cytosol. Cytochrome c promotes the apoptosome formation and initiator caspase 9 activation [10].

Members of the Bcl-2 protein family are critical regulators of apoptosis and include two subgroups of proteins. Pro-survival members, Bcl-2, Bcl-x1, Mcl-1, Bcl-w and A1, have anti-apoptotic activity, while pro-apoptotic proteins, Bax, Bak, Bok, Bid, Bad, Bim, Bik, Blk, Hrk, Noxa, Puma, promote apoptosis [16, 30].

Extrinsic and intrinsic apoptotic pathways merge into the activation point of effector caspases, such as caspase 3, caspase 6 and caspase 7, which directly activate apoptotic cell death [11, 28]. There are several anti-apoptotic proteins over-expressed and on the other hand, some pro-apoptotic proteins are non-functional or low expressed during apoptosis deregulation in malignant tumor [23]. Changes in gene expression associated with deregulation in apoptotic cell death pathway may cause increased proliferation of tumor cells, their uncontrolled growth [1] and resistance to used drugs [16].

3 Methods for Apoptotic Signaling Pathway Study in Cancer Cells

Identification of new targets associated with proliferation, apoptosis, angiogenesis and invasion of cancer cells may help to improve treatment strategies of this disease. Specific microfluidic gene panel arrays are used for gene expression profiling of key ligands, receptors, intracellular modulators, transcriptional factors associated in programmed cell death regulation. Methods such as qRT-PCR based apoptotic microarrays provide comparisons of changes in specific gene expression levels in tumor and healthy biopsy tissue. Recognizing the differences in the expression of pro-apoptotic and anti-apoptotic genes in tumor/control tissues may revealed in the development of novel treatment strategies [1, 36].

Detection of different levels of apoptotic proteins by Western blot analysis, immunoprecipitation or immunohistochemistry is also very important. Quantification of necrotic cells and apoptotic index (percentage of apoptotic cells or apoptotic bodies in 100 viable tumor cells in tumor tissue sample) are considered important prognostic biomarker in tumor tissue and contribute in prediction of tumor cells response to chemotherapy [12].

4 Apoptosis and Specific Molecular Targeting Therapy in Cancer

Regulation of anti-apoptotic proteins of Bcl-2 family should be one of the possibilities in the affection of tumor cell survival. Several compounds, such as sodium butyrate, valproic acid and histone deacetylase inhibitors, which can decrease expression of

anti-apoptotic proteins have been described in recent studies [17, 18]. Antisense oligonucleotides or short single-stranded synthetic deoxyribonucleotides can bind to specific mRNA region and cause its degradation [2].

4.1 Small Molecule Inhibitors of Bcl-2 Family Proteins

Characteristic features of several solid tumors are increased expression levels of Bcl-2 family anti-apoptotic proteins, or decreased levels of pro-apoptotic proteins which result in a subsequent inhibition of programmed apoptotic cell death [16]. Over the last few years several small BH3 mimetics molecules have been synthesized [5, 9, 21]. Apoptotic inhibitors bind selectively to specific anti-apoptotic proteins, inhibit them and inactivate them. Therefore, these specific inhibitors may be useful in the apoptotic studies [7, 30].

Nowadays, the most potent apoptotic inhibitor ABT-737, and its orally active analog ABT-263, binds to the hydrophobic pocket of anti-apoptotic proteins Bcl-2, Bcl-xl and Bcl-w with nanomolar affinity. The released of pro-apoptotic proteins activate Bax and Bak proteins which induce apoptosis or can make cancer cell more sensitive to cytotoxic agents in a range of tumor types [21, 31, 34]. Tumor cell resistance to ABT-737 is associated with high levels of anti-apoptotic Mcl-1 protein. Mcl-1 may inhibit Bcl-2 family pro-apoptotic proteins and block apoptotic cell death [4, 14, 15]. Therefore, the Mcl-1 protein inhibition is a promising target for tumor therapy and can lead to the inhibition of cell proliferation and tumor metastasis [20]. MIM-1 is a novel molecule which selectively binds to the BH3 groove of Mcl-1, with Bak-dependent apoptotic activity and limited thus cell-type dependent potency [5]. In several studies were detected increased levels of Bcl-xl proteins in solid tumor cells [29, 31]. Degterev et al. designed new small molecular inhibitor BH3I, which selectively binds to anti-apoptotic Bcl-xl protein [9].

Selective apoptotic inhibitors may serve as a prototypes for development of the next generation small molecules that effectively reduce the apoptotic threshold in cancer cells.

4.2 Inhibitors of Apoptosis Proteins

Inhibitors of apoptosis proteins are structurally and functionally similar proteins, which are involved in apoptosis, cytokinesis and signal transduction processes. Thus far, eight IAPs (BIRC1–BIRC8) were identified and categorized by structural homology into three subgroups. All IAPs proteins share common structures of baculovirus IAP repeat (BIR) domains to bind and inactivate caspases [33]. Elevated expression of IAPs, preserves cell from apoptotic stimuli and has been reported in several human malignant tumors [19]. The most efficient caspase inhibitor is BIRC4 (XIAP)

among the IAPs protein family. Apoptosis inhibition by BIRC4 is mainly coordinated through binding to initiator caspase 9 and effector caspase 3 and caspase 7 [11].

Antisense oligonucleotides interact with BIRC4 mRNA and ensure its degradation [6]. Small molecular inhibitors can block BIRC4 protein activity and release effector caspase 3 in cell [25]. The levels of Smac protein, inhibitor of BIRC4, is decreased in solid tumor cells and thus cause their survival [27]. Smac mimetic elements increase the activity of effector caspase 3 and therefore represent potential therapeutic opportunity in solid tumor treatment [3].

5 Conclusion

Apoptosis is a comprehensive mechanism involved multiple cell pathways. In process of oncogenesis, cancer cells display deregulations of some apoptotic pathways. The upregulation of anti-apoptotic or downregulation of pro-apoptotic Bcl-2 family proteins are common features of tumor cells. Solid tumors have higher levels of Bcl-2, Mcl-1, Bcl-xl and IAPs proteins comparing to normal cells. Their potential inhibitors should represent novel experimental compounds, therapeutic approaches and opportunities in cancer treatment. Recognition and understanding of molecular mechanisms and signaling pathways associated with programmed cell death in tumor cells may reveal promising treatment strategies for malignant tumors.

Acknowledgments This work was supported by Ministry of Health of the Slovak Republic under the contract MZ SR 2012/25-UKMA-2, by the project Competence Center for Research and Development in the Field of Diagnostics and Therapy of Oncological Diseases, ITMS: 26220220153 and by the project “Biomedical Center Martin” ITMS code: 26220220187, the project is co-financed from EU sources.

References

1. Blahovcová, E., Richterová, R., Kolarovski, B., Dobrota, D., Račay, P., Hatok, J.: Apoptosis-related gene expression in tumor tissue samples obtained from patients diagnosed with glioblastoma multiforme. *Int. J. Mol. Med.* **36**, 1677–1684 (2015)
2. Caruso, G., Caffo, M.: Antisense oligonucleotides in the treatment of cerebral gliomas. Review of concerning patents. *Recent Pat. CNS Drug Discov.* **1**, 2–12 (2014)
3. Chen, D.J., Huerta, S.: Smac mimetics as next cancer therapeutics. *Anti-Cancer Drugs* **20**, 646–658 (2009)
4. Chen, S., Dai, Y., Harada, H., Dent, P., Grant, S.: Mcl-1 down-regulation potentiates ABT-737 lethality by cooperatively inducing Bak activation and Bax translocation. *Cancer Res.* **67**, 782–791 (2007)
5. Cohen, N.A., Stewart, M.L., Gavathiotism, E., Tepper, J.L., Bruekner, S.R., Koss, B., Opferman, J.T., Walensky, L.D.: A competitive stapled peptide screen identifies a selective small molecule that overcomes MCL-1-dependent leukemia cell survival. *Chem. Biol.* **19**, 1175–1186 (2012)
6. Crooke, S.T.: Molecular mechanism of antisense drugs: RNaseH. *Antisense Nucleic Acid Drug Dev.* **8**, 133–134 (1998)

7. Dalafave, D.S., Prisco, G.: Inhibition of antiapoptotic BCL-XL, BCL-2, and MCL-1 proteins by small molecule mimetics. *Cancer Inform.* **9**, 169–177 (2010)
8. Dean, E.M., Ranson, M., Blackhall, F., Holt, S.V., Dive, C.: Novel therapeutic targets in lung cancer: inhibitor of apoptosis proteins from laboratory to clinic. *Cancer Treat. Rev.* **33**, 203–212 (2007)
9. Degterev, A., Lugovskoy, A., Cardone, M., Mulley, B., Wagner, G., Mitchison, T., Yuan, J.: Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. *Nat. Cell Biol.* **3**, 173–182 (2001)
10. Galluzzi, L., Zamzami, N., de La, Motte Rouge, T., Lemaire, C., Brenner, C., Kroemer, G.: Methods for the assessment of mitochondrial membrane permeabilization in apoptosis. *Apoptosis* **12**, 803–813 (2007)
11. Ghavami, S., Hashemi, M., Ande, S.R., Yeganeh, B., Xiao, W., Eshraghi, M., Bus, C.J., Kadkhoda, K., Wiechec, E., Halayko, A.J., Los, M.: Apoptosis and cancer: mutations within caspase genes. *Am. J. Med. Genet.* **46**, 497–510 (2009)
12. Gkogkou, Ch., Frangia, K., Saif, M.W., Trigidou, R., Syrigos, K.: Necrosis and apoptotic index as prognostic factors in non-small cell lung carcinoma: a review. *Springerplus* **3**, 120 (2014). doi:10.1186/2193-1801-3-120
13. Hanahan, D., Weinberg, R.A.: Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011)
14. Hann, ChL, Daniel, V.C., Sugar, E.A., et al.: Therapeutic efficacy of ABT-737, a selective inhibitor of BCL-2, in small cell lung cancer. *Cancer Res.* **68**, 2321–2328 (2008)
15. Kang, M.H., Reynolds, C.P.: Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. *Clin. Cancer Res.* **15**, 1126–1132 (2009)
16. Kelly, G., Strasser, A.: The role of Bcl-2 and its pro-survival relatives in tumourigenesis and cancer therapy. *Cell Death Differ.* **18**, 1414–1424 (2011)
17. Kim, E.H., Kim, H.S., Kim, S.U., Noh, E.J., Lee, J.S., Choi, K.S.: Sodium butyrate sensitizes human glioma cells to TRAIL-mediated apoptosis through inhibition of Cdc2 and the subsequent downregulation of survivin and XIAP. *Oncogene* **46**, 6877–6889 (2005)
18. Kuendgen, A., Lübbert, M.: Current status of epigenetic treatment in myelodysplastic syndromes. *Ann. Hematol.* **8**, 601–611 (2008)
19. La Casse, E.C., Baird, S., Korneluk, R.G., MacKenzie, A.E.: The inhibitors of apoptosis IAPs and their emerging role in cancer. *Oncogene* **17**, 3247–3259 (1998)
20. Li, R.Y., Chen, LCh., Zhang, H.Y., Du, W.Z., Feng, Y., Wang, H.B., Wen, J.Q., Liu, X., Li, X.F., Sun, Y., Yang, D.B., Jiang, T., Li, Y.L., Jiang, ChL: MiR-139 inhibits Mcl-1 expression and potentiates TMZ-induced apoptosis in glioma. *CNS Neurosci. Ther.* **19**, 477–483 (2013)
21. Oltersdorf, T., Elmore, S.W., Shoemaker, A.R., et al.: An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677–681 (2005)
22. Portt, L., Norman, G., Clapp, C., Greenwood, M., Greenwood, M.T.: Anti-apoptosis and cell survival: a review. *Biochim. Biophys. Acta* **1813**, 238–259 (2011)
23. Račay, P., Jurečeková, J., Štefániková, A., Kliková, K., Hatok, J., Dobrota, D.: Bcl-2 family proteins in development and treatment of malignant diseases. *Acta Medica Martiniana* **11**, 15–23 (2011)
24. Sayers, T.J.: Targeting the extrinsic apoptosis signaling pathway for cancer therapy. *Cancer Immunol. Immunother.* **60**, 1173–1180 (2011)
25. Schimmer, A.D., Dalili, S., Batey, R.A., Riedl, S.J.: Targeting XIAP for the treatment of malignancy. *Cell Death Differ.* **13**, 179–188 (2006)
26. Schmitt, E., Pacquet, C., Beauchemin, M., Bertrand, R.: DNA-damage response network at the crossroads of cell-cycle checkpoints, cellular senescence and apoptosis. *J. Zhejiang Uni. Sci. B.* **8**, 377–397 (2007)
27. Sekimura, A., Konishi, A., Mizuno, K., Kobayashi, Y., Sasaki, H., Yano, M., Fukai, I., Fujii, Y.: Expression of Smac/DIABLO is a novel prognostic marker in lung cancer. *Oncol. Rep.* **11**, 797–802 (2004)
28. Shi, Y.: Mechanisms of caspase activation and activation and inhibition during apoptosis. *Mol. Cell.* **9**, 459–470 (2002)

29. Tahir, S.K., Yang, X., Anderson, M.G., Morgan-Lappe, S.E., Sarthy, A.V., Chen, J., Warner, R.B., Ng, S.C., Fesik, S.W., Elmore, S.W., Rosenberg, S.H., Tse, C.: Influence of Bcl-2 family members on the cellular response of small-cell lung cancer cell lines to ABT-737. *Cancer Res.* **67**, 1176–1183 (2007)
30. Thomas, S., Quinn, B.A., Das, S.K., Dash, R., Emdad, L., Dasgupta, S., Wang, X.Y., Dent, P., Reed, J.C., Pellecchia, M., Sarkar, D., Fisher, P.B.: Targeting the Bcl-2 family for cancer therapy. *Expert Opin. Ther. Targets* **17**, 61–75 (2013)
31. Tse, C., Shoemaker, A.R., Adickes, J., et al.: ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res.* **9**, 3421–3428 (2008)
32. Vaux, D.L.: Caspases and apoptosis—biology and terminology. *Cell Death Differ.* **6**, 493–494 (1999)
33. Verhagen, A.M., Coulson, E.J., Vaux, D.L.: Inhibitor of apoptosis proteins and their relatives: IAPs and other BIRPs. *Genome Biol.* **2**, REVIEWS3009 (2001)
34. Vogler, M.: Targeting BCL2 proteins for the treatment of solid tumours. *Adv. Med.* **2014**, 943648 (2014). doi:[10.1155/2014/943648](https://doi.org/10.1155/2014/943648)
35. Westhoff, M.A., Brühl, O., Nonnenmacher, L., Karpel-Massler, G., Debatin, K.M.: Killing me softly - future challenges in apoptosis research. *Int. J. Mol. Sci.* **15**, 3746–3767 (2014)
36. Zubor, P., Hatok, J., Moricova, P., Kajo, K., Kapustova, I., Mendelova, A., Racay, P., Danko, J.: Gene expression abnormalities in histologically normal breast epithelium from patients with luminal type of breast cancer. *Mol. Biol. Rep.* **42**, 977–988 (2015)

Part III
Modeling and Simulation

Application of Reverse Engineering in Supporting the Treatment of Pectus Carinatum

Magdalena Antonowicz, Anita Kajzer and Wojciech Kajzer

Abstract This paper presents the application of reverse engineering in supporting the design and modelling of the personalised prototype of the orthotic bracing in the treatment of pectus carinatum. 3D scans of the normal thoracic cage were used, which were necessary to simulate pectus carinatum. Based on the scans, the skeletal system of the pectus carinatum and the normal skeletal system were designed. In summary, it can be stated that it was possible to design the personalised external stabiliser and produce an animation of the principle of its operation.

Keywords 3D scan · Pectus carinatum · Pectus bracing · Reverse engineering

1 Introduction

The dynamic development of reverse engineering helps to transfer a physical external object to a computer model, which enables the shortening of the design process and tests the simulation of the functional properties, supporting and helping make prototypes and completed products. The stages of rapid prototyping technology are based on the idea that it is necessary to generate an expected model (Fig. 1), and the whole process is composed of a few essential levels, with the first being the 3D scanning process [13].

The characteristic feature of the surface digitalisation of the object is, among others, the analogical-digital transformation directly linked with object's information processing. This is why one of the most prospective applications of reverse engineering is supporting treatment and using it in the entire field of biomedical engineering as a tool in the operational planning process. Moreover, it is also used to design medical devices that support the treatment of diseases, e.g. in orthopaedics, thoracoscopy, traumatology and paediatric surgery [7, 9].

M. Antonowicz (✉) · A. Kajzer · W. Kajzer
Faculty of Biomedical Engineering, Department of Biomaterials and Medical Devices
Engineering, Silesian University of Technology, Roosevelta 40, Zabrze, Poland
e-mail: magdalena.antonowicz@polsl.pl

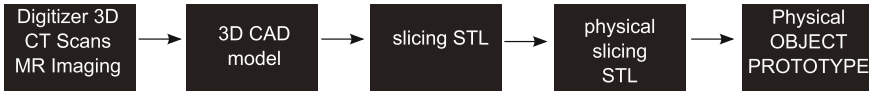


Fig. 1 The idea of rapid prototyping method [13]

In the literature, cases of congenital deformations of the chest wall in children, such as pectus excavatum and carinatum forms, are often described [5, 6, 8, 15, 17]. Pectus carinatum is the deformity of the skeletal system related to the front chest wall. The defect is characterised by sternum and costal cartilage deformation, which leads to anterior protrusion. The most important and probably the only treatment of the above disease is to handle the psychological and aesthetical aspects [5]. Further, there are three main types of pectus carinatum: symmetrical, asymmetrical, and mixed (with various intensities and areas of occurrence) [8]. It is said that this deformity is more likely to affect men than women. It is essential to select the individual treatment method to cure the pectus carinatum. The application of the chosen treatment method is directly related to the type, level of deformation, anatomical form of the chest and physiological patient abilities [6]. To define the size of the deformity, it is necessary to do CT scans of the chest.

Increasingly, to avoid radiation of the patient and to monitor the steps of treatment, the medical imaging is supported by 3D scanning images [17]. High-intensity deformities of the front chest wall, in most cases, are treated with minimally invasive repair methods with the use of correctly profiled implants that help to restore the anatomical shape of the chest [10, 11]. In this case, it is necessary to use surgery tools with suitable properties that ensure a safe implantation process, mainly in terms of thermal consideration. Further, in the event of a patient being qualified to a group with light or moderate imperfections or after the surgical intervention of pectus carinatum, it is recommended to use external bracings for treatment [4]. The implants are made of steel Cr-Ni-Mo [11], but, for patients who are allergic to components of this alloy, titanium alloys are used [1–3, 12, 20]. The basic principle of the implants is to provide constant pressure for a period of time on the most convex area, causing the proper modelling of cartilages. With the use of external bracings, we can avoid post-operative complications, scars and pain; moreover, the use of a brace is an effective way to correct pectus carinatum in a visible way [6]. Through computer methods, it is possible to create a personalised stabiliser construct to cure pectus carinatum, which is adjusted to individual anthropometric patient features.

The aim of the present work was to elaborate on the multistage procedure of making a personalised prototype of an external stabiliser for the treatment of pectus carinatum supported by reverse engineering techniques. Based on 3D scans of a normal male chest, the pectus carinatum model, with an intensity that is, according to the literature, the most common (type 5 according to the Willital scale—symmetrical deformation) was created [16]. Thus, it was possible to design an external stabiliser model to cure this type of imperfection.

2 Normal Thoracic Cage and Pectus Carinatum 3D Models

The carried out biomedical reconstructions are more often used to create instructional, reconstructive models, and implant prototypes or rehabilitation devices [7]. Reverse engineering is essential in beginning procedures leading to the design of personalised external stabilisers.

The performed stages that lead to the precise design of personalised stabiliser prototypes began with transferring the real 3D shape of the normal chest to the digital form with the usage of a handy digitiser (HandyScan 3D RevScan). Firstly, the test stand was composed of a 3D RevScan, a computational unit with software VxScan and a patient (male, age 24). On the chest surface, the marks were placed carefully, as they influence the quality and accuracy of the results and enable scans to be put together (Fig. 2).

To obtain the full image of the male chest, it was necessary to conduct two tests of the front part of the chest and one of the back part (Fig. 3). Scanning took about 30 minutes.

The dataset from the chest surface allowed to obtain a 3D point cloud, in which every point corresponded to a point on the real chest (Fig. 4).

The next stage of the digital transformation of the chest surface was to filter and edit the point cloud to obtain consistent and optimised points made by GeoMagic

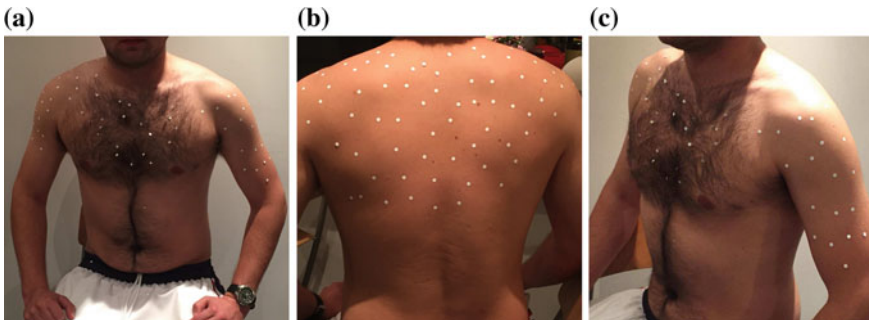


Fig. 2 Human body models with positioning targets: **a** frontal view, **b** rear view, **c** three-dimensional view

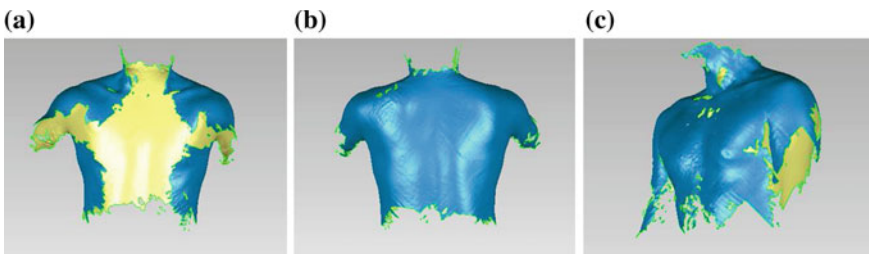


Fig. 3 Chest with normal structure scans 3D: **a** frontal view, **b** rear view, **c** three-dimensional view

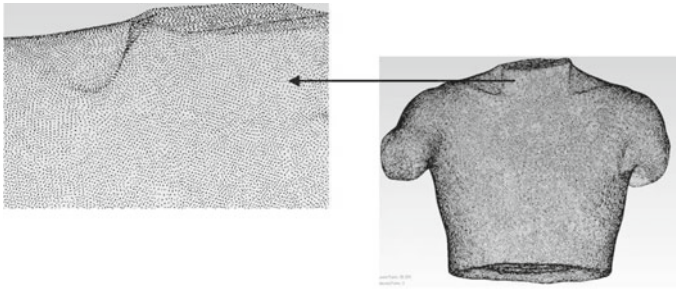


Fig. 4 Points cloud of scanned chest

Studio 2012 based on the STL (STereoLithography) files. The filtration action was done via the removal of the noise measurements and by combining and optimising components of directional point clouds obtained from particular measurements. A point cloud, as a 3D mapping of the chest surface, is not the model form that is suitable for direct use in CAD computer systems where the triangulation of point clouds was carried out. The point cloud showed the triangulation surface establishment with the Delanuy Method (Fig. 5a) and the Laplace Surface Smoothing Method, which are based on the translation of internal triangulation nodes. New point locations were calculated according to following formula [18, 19]

$$x = \frac{1}{n} \sum_{i=1}^n x_i \tag{1}$$

where n—number of points edge-connected with given point.

Then, the normal chest geometry was used in CAD systems (in this case, Autodesk Inventor Professional 2015—(Fig. 5b).

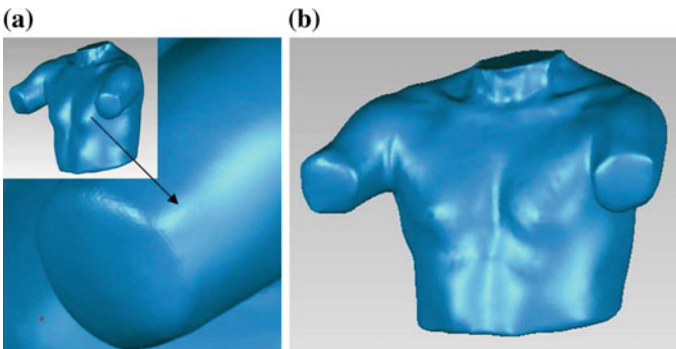


Fig. 5 Transformation of the model: **a** triangle mesh with shadows, **b** Laplace'a smoothing

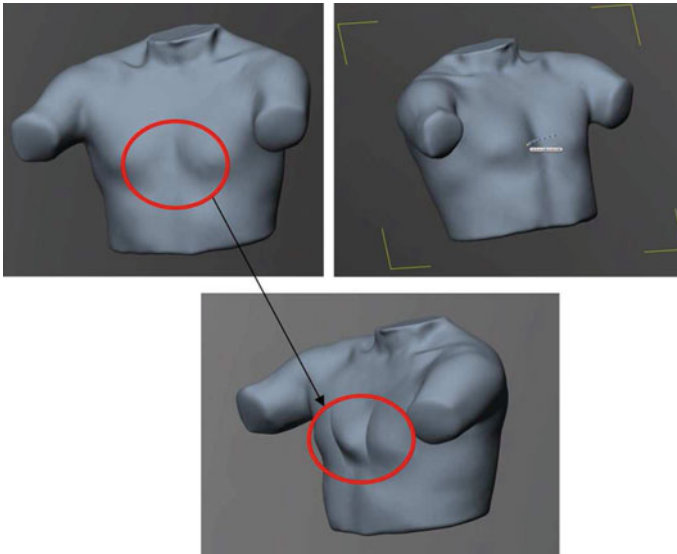


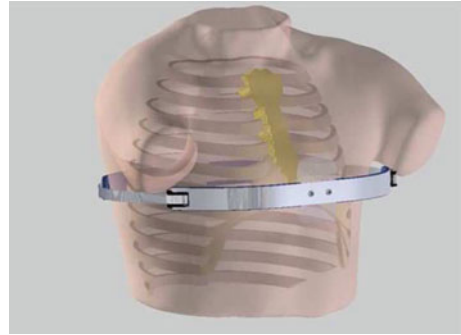
Fig. 6 Pectus carinatum simulation

The indispensable element in designing an external personalised stabiliser prototype was the pectus carinatum model. Hence, on the basis of the generated 3D normal thoracic cage, there was a model that simulated pectus carinatum with the use of Inventor Fusion 2013 software. Based on the literature [16], patients with pectus carinatum are exposed to symmetrical sternum grip deformation; therefore, the simulation was done in this area. Displacement of the sternum was performed by means of interpolational plane curve surface modelling, which consisted of points that were altered to obtain a 30 mm sternum displacement in regard to the normal thoracic cage (Fig. 6).

3 Simplified Geometrical Model of Thoracic Cage

On the grounds of 3D models of the surface of the normal thoracic cage and pectus carinatum, a simplified 3D model of a thoracic cage was made with the use of Inventor Professional 2015 software. This was done to design the external stabiliser in a precise way. As part of the interaction principle of an external stabiliser to the bone arrangement of the thoracic cage, there were four thoracic cage models designed with different sternum deformations ranging from of 0–30 mm. The elaborated model of the bone arrangement of pectus carinatum enabled to design a personalised external stabiliser to cure this type of imperfection. Further, the surface 3D thoracic cage model was put on bone structures, then, based on that model, the dimensions and shape of the external stabiliser were mapped. The design began via preliminary

Fig. 7 Main elements of external stabilizer



construction, utility and aesthetic assumptions. Construction elements counteracting damage, irritation and protecting the skin surface were taken into consideration. The external stabiliser model was created in Inventor Professional 2015 software, which contained 13 permanent and two additional elements. All parts were designed separately and connected to each other at the end. The main and most important elements were the bars, which were positioned on the anterior and the posterior chest, and connected to each other with a strap ratchet kit and buckles to enable accurate control according to the chest (Fig. 7). Moreover, on the middle part of the curve, an element compressing the biggest deformation was placed. The protection foams were also taken into consideration.

To provide a pleasing aesthetic appearance of the orthosis and to cover stabiliser elements, a case was proposed. The proper position of the stabiliser was assured via adequate compression by ratchet strips, but, to increase stabilisation, additional suspenders were designed (Fig. 8).

The last step was to merge the simulated pectus carinatum model skeletal system with different sternum locations and to design the pectus carinatum model to show the behaviour of the skeletal system under stabiliser compression. This was

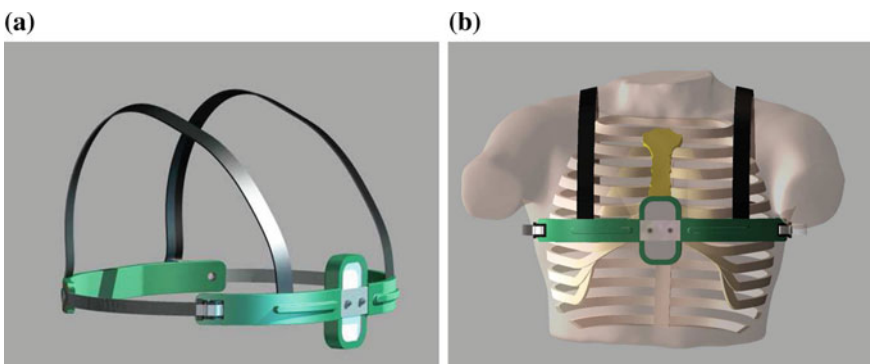


Fig. 8 An external stabilizer model for curing pectus carinatum: **a** construction, **b** placed on patient

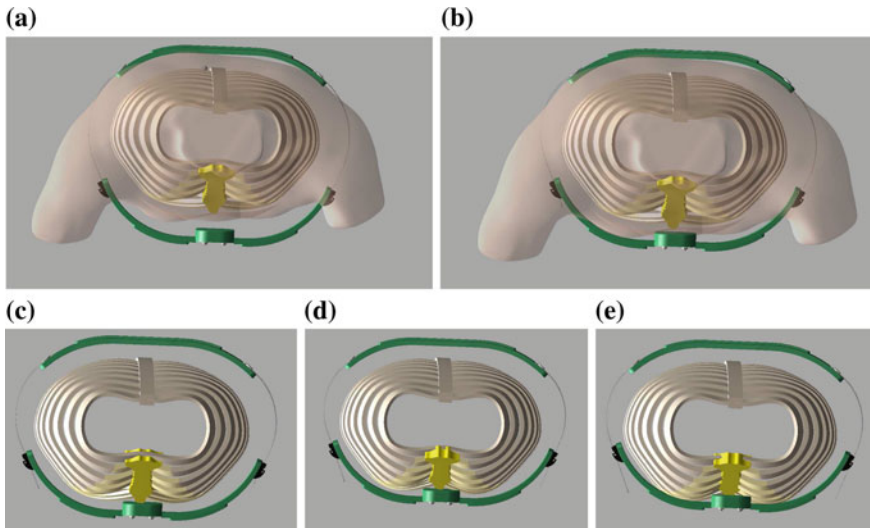


Fig. 9 Stages of external stabilizer action

performed via Inventor Professional and Inventor Studio environment software. In the animation, there were five stages of stabiliser action shown:

- Stage I—placing the stabiliser on the pectus carinatum—(Fig. 9a),
- Stage II—compression number one, change of sternum position—(Fig. 9b),
- Stage III—compression number two, change of sternum position—(Fig. 9c),
- Stage IV—compression number three, change of sternum position—(Fig. 9d),
- Stage V—sternum arrangement in proper position—(Fig. 9e).

The animation allowed to present the behaviour of the external stabiliser in regard to the chest and sternum by correcting its position. The procedure shown in Fig. 10, which begins with scanning and leads to obtaining the model, is the base to begin making a physical prototype of the stabiliser model; it enables a personalised fitted construction to patient’s imperfection.

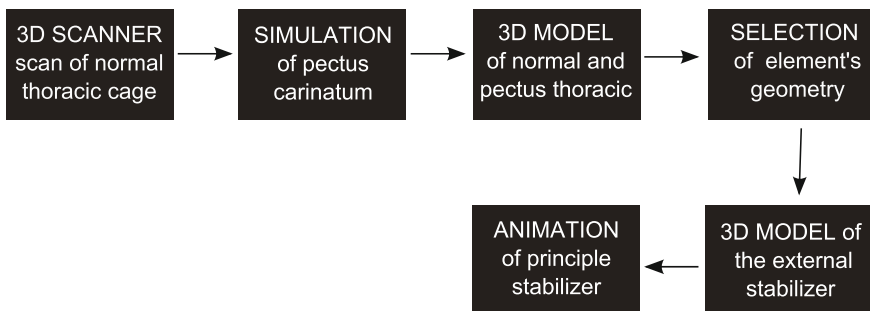


Fig. 10 The personalized external stabilizer design procedures

4 Conclusion

Reverse engineering was used to image the external surface of the thoracic cage in preliminary thoracic cage size determination. On the other hand, in the event of illustrating a patient's internal structure and viscus, the essential tool for medical illustrating was the CAT scan. Obtained image projections allowed to determine the level and type of thoracic cage deformation, symmetry, rotation of sternum, and hearth location. It was also possible to rate the anatomical disturbances expressed by the Haller indicator [14].

The pectus carinatum imperfection causes mental discomfort and an abnormal Haller indicator, which is why it is essential to find effective and rapid curing methods. Further, reverse engineering and CAD software played a useful role in this project; they supported the medical imaging techniques that avoided exposing the patient to radiation during CAT scanning and helped to design an external stabiliser in a rapid and precise way. However, based on particular 3D scanning images, it is possible to fit and prepare a construction project depending on a patient's defect.

In the present paper, the procedures showing the preparation of personalised male external stabiliser design methods were introduced. Based on obtained 3D thoracic cage images, it was possible to simulate pectus carinatum and to design the geometry of an external stabiliser in an accurate way. To visualise the behaviour and principles of the operation of the stabiliser in accordance to the thoracic cage, an animation was prepared. This assessment is the first and foremost action that leads to physical prototype stabiliser model accomplishment.

The difficult part when beginning construction was the individual anatomical patient conditioning, type and size of the defect. This is why the 3D images became indispensable elements to begin the design processes. This paper can be expanded by adding construction materials for the external stabiliser and by performing bearing capacity tests. To verify the stabiliser's construction, it is also possible to conduct the finite element analysis by distinguishing plastic cartilages in a real skeletal thoracic cage model. All the stages might lead to the manufacturing of the external stabiliser prototype for pectus carinatum treatment. In conclusion, approaches to this type of issue were never before presented in the literature.

References

1. Basiaga, M., Paszenda, Z., Walke, W., Karasiński, P., Marciniak, J.: Electrochemical Impedance Spectroscopy and corrosion resistance of SiO₂ coated cpTi and Ti-6Al-7Nb alloy. *Information Technologies in Biomedicine. Advances in Intelligent Systems and Computing*, vol. 284, pp. 411–420. Springer (2014)
2. Basiaga, M., Walke, W., Paszenda, Z., Kajzer, A.: The effect of EO and steam sterilization on the mechanical and electrochemical properties of titanium Grade 4. *Materiali in tehnologije/Mater. Technol.* **50**(1), 153–158 (2016)

3. Basiaga, M., Walke, W., Paszenda, Z., Karasiński, P., Szewczenko, J.: The effects of a SiO₂ coating on the corrosion parameters cpTi and Ti-6Al-7Nb alloy. *Biomater* **4**, 1 (2014). doi:[10.4161/biom.28535](https://doi.org/10.4161/biom.28535)
4. Cohee, A.S., Lin, J.R., Frantz, F.W., Kelly, R.E.: Staged management of pectus carinatum. *J. Pediatr. Surg.* **48**, 315–320 (2013)
5. Czernik, J.: Powikłania w chirurgii dziecięcej (Complications in Children's Surgery). Wydawnictwo Lekarskie PZWL, Warszawa (2009) (in Polish)
6. Desmarais, T.J., Keller, M.S.: Pectus carinatum. *Curr. Opin. Surg.* **25**, 375–379 (2013)
7. Dybała, B.: Integracja i spójność modeli w inżynierii odwrotnej: Wybrane aspekty technicznych i medycznych zastosowań Reverse Engineering (Integration and Cohesion of Models in Reverse Engineering: Selected Aspects of Technical and Medical Applications of Reverse Engineering). Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław (2013). (in Polish)
8. Fokin, A.A., Steuerwald, N.M., Ahrens, W.A., Allen, K.E.: Anatomical, Histologic, and Genetic Characteristics of Congenital Chest Wall Deformities. *Thoracic and Cardiovascular Surgery*, pp. 44–57. Elsevier (2009)
9. Glinkowski, W., Michonski, J., Sitnik, R., Witkowski, M.: 3D diagnostic system for anatomical structures detection based on a parameterized method of body surface analysis. *Information Technologies in Biomedicine: Advances in Intelligent and Soft Computing*, vol. 69, pp. 153–164. Springer (2010)
10. Kajzer, W., Kajzer, A., Gzik-Zroska, B., Wolański, W., Janicka, I., Dzielicki, J.: Comparison of numerical and experimental analysis of plates used in treatment of anterior surface deformity of chest. In: Piętka, E., Kawa, J. (eds.): *ITIB (2012)*, LNCS, vol. 7339, pp. 319–330. Springer, Berlin (2012)
11. Kajzer, A., Kajzer, W., Dzielicki, J., Matejczyk, D.: The study of physicochemical properties of stabilizing plates removed from the body after treatment of pectus excavatum. *Acta Bioeng. Biomech.* **2**, 35–44 (2015). doi:[10.5277/ABB-00140-2014-02](https://doi.org/10.5277/ABB-00140-2014-02)
12. Marciniak, J., Szewczenko, J., Kajzer, W.: Surface modification of implants for bone surgery. *Arch. Metallur. Mater.* **60**(3B), 13–19 (2015)
13. Miecielica, M.: Rapid prototyping—metody i możliwości zastosowania w inżynierii biomedycznej (Rapid Prototyping—Methods and Possibilities of Application in Biomedical Engineering). AGH, Kraków (2009) (in Polish)
14. Park, C.H., Kim, T.: Seok, J.H., Jeon, I., Lee, S.: The etiology of pectus carinatum involves overgrowth of costal cartilage and undergrowth of ribs. *J. Pediatr. Surg.* **49**, 1252–1258 (2014)
15. Robicsek, F.: Surgical repair of pectus excavatum and carinatum. *Semin. Thorac. Cardiovasc. Surg.*, 64–75 (2009)
16. Stoba, Cz., Willitala, G.H., Sołtysiak, P.K.: Atlas chirurgii dziecięcej (Atlas of Children's Surgery). Bernardinum, Pelplin, pp. 31–41 (2008) (in Polish)
17. Wong, K.E., Gorton, G.E., Tashjian, D.B., Tirabassi, M.V., Moriarty, K.P.: Evaluation of the treatment of pectus carinatum with compressive orthotic bracing using three dimensional body scans. *J. Pediatr. Surg.* **49**(6), 924–927 (2014)
18. Wyleżół, M.: Metodyka modelowania na potrzeby inżynierii rekonstrukcyjnej (Modelling Methodology for the Purpose of Reverse Engineering). Wydawnictwo Politechniki Śląskiej, Gliwice (2013) (in Polish)
19. Wyleżół, M., Muzalewska, M.: Metodyka modelowania w inżynierii biomedycznej z użyciem inżynierii rekonstrukcyjnej (Modelling Methodology in Biomedical Engineering with the Use of Reverse Engineering), *Mechanik 02/2015* (2015) (in Polish)
20. Ziębowicz, A., Ziębowicz, B., Bączkowski, B.: Electrochemical behavior of materials used in dental implantological systems. *Solid State Phenom.* **227**, 447–450 (2015)

Breast Deformation Modeling Based on MRI Images, Preliminary Results

Marta Danch-Wierzchowska, Damian Borys and Andrzej Swierniak

Abstract Increasing number of people are suffering from cancer these days. The highest incidence among women is breast cancer. One of the main sources of diagnostic information is Magnetic Resonance Imaging (MRI). Another one is Positron Emission Tomography with Computer Tomography (PET-CT). Both examinations take place in different patient positions (prone and supine respectively). The only way to obtain complete diagnostic information is to bring MRI images into PET-CT space and compare both. Our preliminary studies focus on creating a simple, deformable finite element model of the breast. Proposed algorithm allows to obtain breast model in supine position, created from MRI images obtained in prone position.

Keywords Breast deformation · Finite element modeling · MRI images

1 Introduction

According to recent studies, breast cancer is the most frequent women's cancer around the world. To help growing number of patients, diagnostics need to be precise and fast, whereas treatment needs to be as personalised as possible.

One of the main sources of diagnostic information used in surgical planning is Magnetic Resonance Imaging (MRI). MRI is used to obtain information about tissue condition, that cannot be obtained by other popular imaging techniques [2]. Another one is Positron Emission Tomography with Computer Tomography (PET-CT). PET data contain informations about tissues metabolics, but their images contain only pixel clouds. PET-CT fusion allows to locate metabolic PET clouds inside the body, on CT scans. Unfortunately, CT scans are less detailed than MRI data. MRI and PET images fusion could help with fast and precise diagnosis. Fused prone PET

M. Danch-Wierzchowska (✉) · D. Borys · A. Swierniak
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Automatic Control, Silesian University of Technology,
Akademicka 16, 44-100 Gliwice, Poland
e-mail: marta.danch-wierzchowska@polsl.pl

and MR images allow to observe both tissue structure and its metabolic activity simultaneously, which increase significantly diagnostic confidence, far more than side-by-side review [8]. The main problem which occurs is that MRI images are obtained in prone patient position with breasts placed in signal enhancement coils, while PET-CT examination is performed in supine position. The breasts in both situations have different shapes, and tissues of interest are differently placed on images. The bigger the breast is, the more its shape differs in both positions [12]. Due to that, a need for easy MRI images transformation into PET-CT space arises. This cannot be obtained by placing patient supine in MRI, or prone during PET-CT, which needs to be used in surgery planning (supine). The only way to achieve this goal is computer image processing. A similar problem was addressed in [1], where mammograms compression was simulated based on simple breast model. MRI image deformation was presented in [5], however a model of breast tissues was very complicated. Recent studies [7, 14] present models that consist of four structures: fat, glandular, cancerous and skin tissues. Building a model for each patient and complicating the structure considered could extend severely computation time, resulting in increasing time of diagnosis.

Yet another issue in model creation is proper tissue segmentation. Segmentation of breast MRI images is still a difficult job, because of a complex structures, field inhomogeneity and image noise, which need to be overcome to create automatic finite element breast models [9]. While segmentation of the breast shape from background is simple enough [10, 11, 15], the main problem is segmentation of a back side of chestwall, which is nearly invisible on MRI images, due to field inhomogeneity caused by signal enhancement coils. To preserve mechanical dependencies in the body, the first idea is to create full body model. Due to invisibility of most of the back side of the body, it is almost impossible to properly segment it. However, our model overcomes difficulty with back side of chestwall segmentation, preserving simultaneously main properties of deformed tissues.

In this work we present a basic idea of a breast model and its deformation, paying a special attention to further usage in clinical practice. Deformation algorithm is based on well-known Finite Element Methods (FEM) best practices. Our preliminary study consists of creating a model based directly on medical images. The novelty of our approach is implementing large breasts deformation, with mesh created directly on medical images, without using any third-party software. A deformation model based on the Lagrange equations of motion [1] was implemented in Matlab and breast mesh was deformed using gravitational forces.

2 Materials and Methods

2.1 Data Acquisition

Breast MRI data acquisition MRI is widely used in medical diagnosis, in particular breast imaging [2]. MRI scans T1-weighted were acquired at the Center of

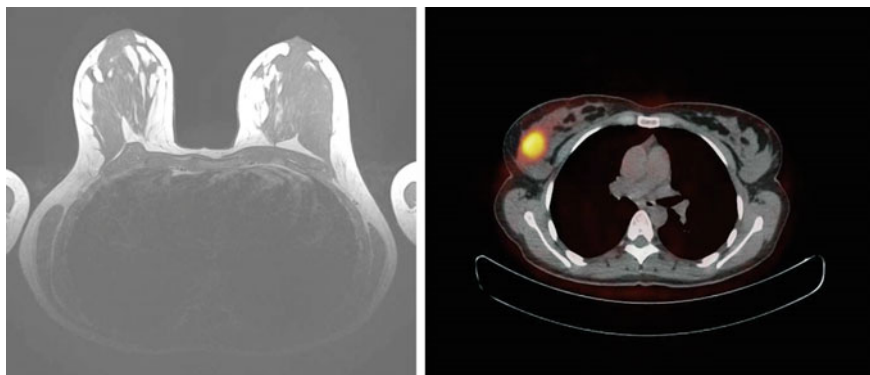


Fig. 1 Exemplary images: on the *left* MRI in prone position and on the *right* PET-CT examination in supine position

Oncology—Maria Skłodowska-Curie Memorial Institute, Branch in Gliwice, Poland using a Siemens scanner. The data consist of 50 axial slices of patient in prone position covering the area of interest. Obtained data cover approximately $0.7 \times 0.7 \times 3$ mm real volume per voxel. Exemplary image is shown in Fig. 1 (left).

Breast PET-CT data acquisition Positron Emission Tomography with Computer Tomography (PET-CT) is fusion from 3D X-ray structural examination (CT) and based on gamma-rays metabolic examination (PET). PET is performed using fluoro-deoxyglucose (FDG), an analogue of glucose to indicate tissue metabolic activity. PET-CT scans were acquired at the Center of Oncology in Gliwice using a Siemens Biograph scanner. The data consist of 289 axial slices of patient in supine position covering almost the whole patient body. Obtained data cover approximately $1.17 \times 1.17 \times 3$ mm real volume per voxel. PET-CT scans are obtained in supine patient position and were used as reference to deformation validation. An exemplary image is shown in Fig. 1 (right).

2.2 Deformation Algorithm

Deformation algorithm consisting of a few main steps based on FEM best practices is shown in the form of block diagram (Fig. 2).

ROI segmentation To obtain breast mask (Fig. 3 (middle)) basing on MRI image (Fig. 3 (left)), a fuzzy c-means algorithm is used [17]. To simplify computation chestwall is modeled as half ellipse, not segmented directly from image.

Dividing into left/right part To speed up computational time, the model is divided into left and right breast. Since there is no difference in methodology we present only one half of the model.

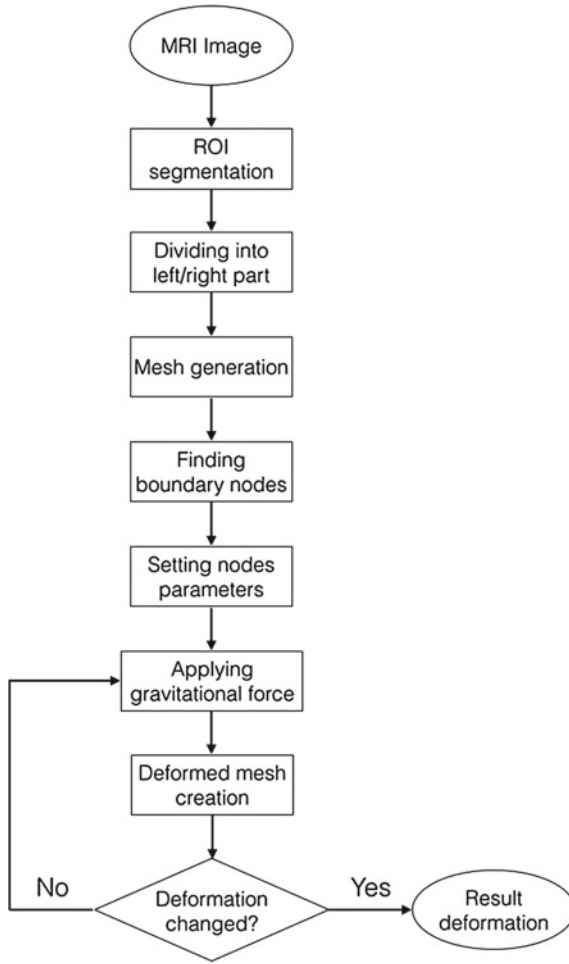


Fig. 2 Block diagram of used deformation algorithm

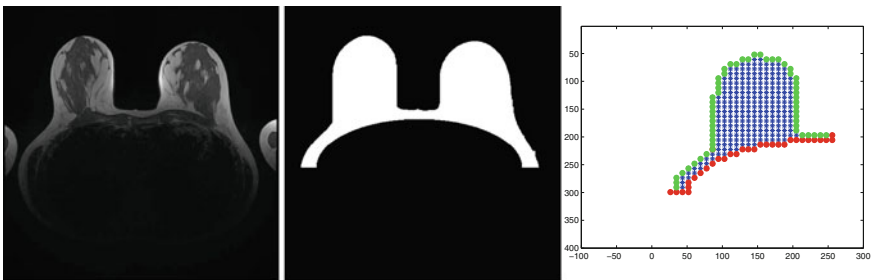


Fig. 3 Exemplary mesh creation stages, from *left* segmented MRI image, segmentation mask, *left* breast mesh (*green*—skin nodes, *red*—stiffened nodes, *blue*—other nodes)

Mesh generation Based on ROI mask, orthogonal mesh is created (Fig. 3 (right)). Distances between nodes depend on image size, here 1 node per 5 image pixels. Due to limits of commonly used clinical hardware, algorithm requires less than 1000 nodes per object (left or right breast).

Finding boundary nodes To set different parameters, the created model is divided into three parts (Fig. 3 (right)): skin nodes—green, chestwall nodes—red and other—blue.

Setting nodes parameters To implement deformation model [1] Lagrange equation of motion was used:

$$\mathbf{M} \frac{\partial^2 \mathbf{q}}{\partial t^2} + \mathbf{D} \frac{\partial \mathbf{q}}{\partial t} + \mathbf{K} \mathbf{q} = \mathbf{g}_q + \mathbf{f}_q, \quad (1)$$

where \mathbf{M} , \mathbf{D} and \mathbf{K} are the mass, damping and stiffness matrices, respectively. The vector \mathbf{q} is displacement along x, y, z axis. \mathbf{g}_q and \mathbf{f}_q are internal and external forces, respectively. Tissue parameters were set according to [4] as one isotropic, homogeneous and incompressible material, described by the neo-Hookean hyperelastic model:

$$W = C_1(I_1 - 3), \quad (2)$$

where C_1 is a material constant, and I_1 is the first invariant of the right Cauchy-Green deformation tensor.

Simplification of the model was reached by taking into account that fat material parameters have much more influence on deformation than fibroglandular ones [9]. Skin nodes constrain collision detection and reaction. Chestwall nodes are set as stationary.

Application of gravitational force Gravitational force is applied to every node (except chestwall nodes) in the mesh. Displacement of each node is calculated according to implemented equation of motion. The procedure is repeated until breast reaches steady state, i.e.: no more displacement is detected.

3 Results

The deformation result is represented as breast mesh in the PET-CT plan (supine). Figure 4 shows selected mesh stages for exemplary image (Fig. 3). The upper left image is original mesh, while the bottom right image contains the resulting mesh. The resulting mesh was compared with corresponding PET-CT image (See Fig. 5). Differences are easily visible. Chestwall shape is not-identical, probably caused by different muscle tension in prone and supine position. However, main shape is preserved. Comparison of PET-CT image with MRI image deformed with created mesh would be easier than comparison PET-CT with original prone MRI image.

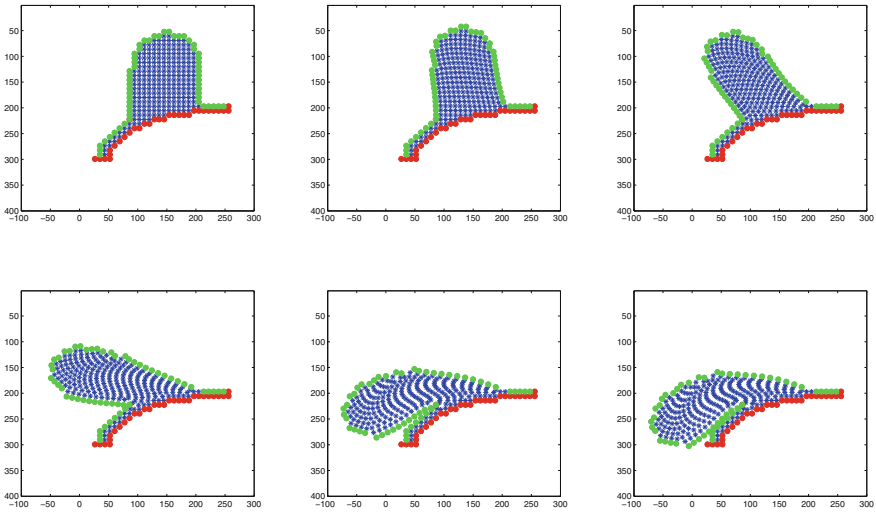


Fig. 4 Selected mesh deformation stages: original mesh (*upper left*), resulting mesh (*bottom right*), (*green*—skin nodes, *red*—stiffened nodes, *blue*—other nodes)

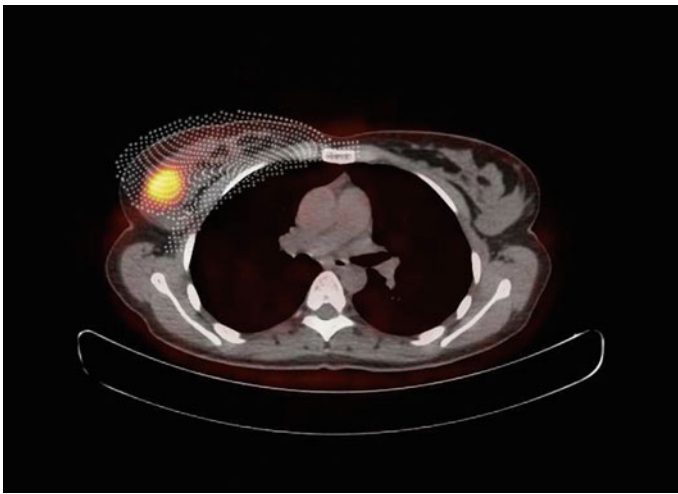


Fig. 5 Comparison of resulting mesh with corresponding PET-CT image

Presented results are obtained in less than 3 min per image on average portable computer, with fully automated algorithm. This allows us to believe, that our algorithm, extended and improved, would be useful in clinical practice.

4 Conclusions and Further Work

It is clear that the results obtained could not be precise. There are many more conditions influencing tissue movement than gravitational force. There is still muscle tension, which influences breast shape. It is impossible to mimic all body dependencies and movements personalized for every patient in limited time. However, there are image processing algorithms, like free-form deformation (FFD), which were successfully used for small non-linear deformation [13]. FFD algorithm could help with precise fitting of MRI image to reference CT image, but is incapable of large image deformation and cannot be used only by itself.

Breast deformation modeling is mainly focused on gravitational force and its impact on the breast shape. One of the main challenges is a compromise between model accuracy and work efficiency. Too complex tissue models could never be used in clinical practice, since adapting them to each patient would be too time-consuming. Developing an automated tool for simplified shell model creation from medical images could speed up deformation analysis significantly. Analysis using available commercial software (i.e. ANSYS, Abacus) [3, 6, 14, 16] gives acceptable results but implementing the model inside it is highly time-consuming. The best solution seems to build a model based directly on medical images, without using any third party software.

The inaccuracy of the model is a result of one image deformation, a 3D model could also mimic breast movement in the anteroposterior axis and maintain continuity of tissues in every direction. Creating MRI images from obtained deformed mesh and proper segmentation methods will be under further investigation of our work.

Acknowledgments This work was supported by the National Science Centre NCN under Grant No. 2011/03/B/ST6/04384 (AS), the National Centre for Research and Development under grant number PBS/32/RJP8/2015/514 (DB) and the Institute of Automatic Control under Grant No. BKM-514/RAU1/2015/t.12 (MDW). This work was supported by Scholarship Program “DoktoRIS—Program stypendialny na rzecz innowacyjnego Slaska”.

References

1. Azar, F.S., Metaxas, D.N., Schnall, M.D.: Methods for modelling predicting mechanical deformations of the breast under external perturbations. *Med. Image Anal.* **6**, 1–27 (2002)
2. Behrens, S., Laue, H., Althaus, M., Boehler, T., Kuemmerlen, B., Hahn, H.K., Peitgen, H.O.: Computer assistance for MR based diagnosis of breast cancer: present and future challenges. *Comput. Med. Imaging Graph.* **31**(4–5), 236–247 (2007)
3. Han, L., Hipwell, J., Taylor, Z., Tanner, C., Ourselin, S., Hawkes, D.J.: Fast deformation simulation of breasts using GPU-based dynamic explicit finite element method. *Digit Mammo.* **6136**, 728–735 (2010)
4. Han, L., Hipwell, J., Mertzaniidou, T., Carter, T., Modat, M., Ourselin, S., Hawkes, D.: A hybrid FEM-based method for aligning prone and supine images for image guided breast surgery. From Nano to Macro: 2011 IEEE International Symposium on Biomedical Imaging, pp. 1239–1242 (2011)

5. Han, L., Hipwell, J.H., Eiben, B., Barratt, D., Modat, M., Ourselin, S., Hawkes, D.J.: A nonlinear biomechanical model based registration method for aligning prone and supine MR breast images. *IEEE Trans. Med. Imaging*. **33**(3), 682–694 (2014)
6. Hopp, T., Dietzel, M., Baltzer, P.A., Kreisel, P., Kaiser, W.A., Gemmeke, H., Ruitter, N.V.: Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization. *Med. Image Anal.* **17**(2), 209–218 (2012)
7. Lee, A.W., Rajagopal, V., Babarenda Gamage, T.P., Doyle, A.J., Nielsen, P.B., Nash, M.P.: Breast lesion co-localisation between X-ray and MR images using finite element modelling. *Med. Image Anal.* **17**(8), 1256–1264 (2013)
8. Moy, M., Noz, M.E., Maguire, G.Q., Melsaether, A., Deans, A.E., Murphy-Walcott, A.D., Ponzo, F.: Role of fusion prone FDG-PET and Magnetic Resonance Imaging of the breasts in the evaluation of breast cancer. *Breast J.* **16**(4), 369–376 (2010)
9. Pathmanathan, P., Gavaghan, D.J., Whiteley, J.P., Chapman, S.J., Brady, J.M.: Predicting tumor location by modeling the deformation of the breast. *IEEE Trans. Biomed. Eng.* **55**(10), 2471–2480 (2008)
10. Song, H., Cui, X., Sun, F.: Breast tissue 3D segmentation and visualization on MRI. *Int. J. Biomed. Im.* 859746 (2013)
11. Srikantha, A., Harz, M.T., Newstead, G., Wang, L., Platel, B., Hegenscheid, K., Mann, R.M., Hahn, H.K., Peitgen, H.-O.: Symmetry-based detection and diagnosis of DCIS in breast MRI. *Med. Imaging 2013: Comput.-Aided Diagn.* 86701E (2013)
12. Ramiao, N., Martins, P., Fernandes, A.A.: Biomechanical properties of breast tissue. *ENBENG 2013*, 1–6 (2013)
13. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* **18**(8), 712–21 (1999)
14. Tanner, C., Hipwell, J.H., Hawkes, D.J., Szekely, G.: Breast shapes on real and simulated mammograms. *Digit Mammo.* **6136**, 540–547 (2010)
15. Wang, L., Filippatos, K., Friman, O., Hahn, H.K.: Fully automated segmentation of the pectoralis muscle boundary in breast MR images. *Med. Imaging 2011: Comput.-Aided Diagn.* 796309 (2011)
16. Wessel, C., Schnabel, J.A., Brady, M.: Realistic biomechanical model of a cancerous breast for the registration of prone to supine deformations. *EMBC 2013*, 7249–7252 (2013)
17. Wu, S., Weinstein, S.P., Conant, E.F., Schnall, M.D., Kontos, D.: Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images. *Med. Phys.* **40**(4), 042301 (2013)

ToF-Data-Based Modelling of Skin Surface Deformation

Jan Juszczyk, Joanna Czajkowska, Bartłomiej Pycinski
and Ewa Piętka

Abstract Nowadays, the imaging techniques followed by advanced image processing algorithms become an indispensable part of diagnostic and surgical procedures. During recent years, an increasing demand for intra-surgical modalities can be noticed. The work aims at modelling the deformation of human skin surface caused by a navigated stick based on the point cloud acquired by a Time-of-Flight (ToF) camera. The data acquired by ToF and optical tracker are synchronized. Then, the skin deformation is modelled by applying a physics engine. The model evaluation is based on a Hausdorff distance. The results prove the applicability of the developed workflow.

Keywords Surface deformation · Physical features · Optical tracker · ToF camera

1 Introduction

Since the 3D imaging techniques had been developed, plenty of their applications in medical field were found, e.g. facial and dental imaging, internal organ measurement [1], surgical planning in orthopaedics [20] or plastic surgery [5]. Fast and accurate 3D scene analysis is now crucial for development of computer aided diagnosis and therapy. New modalities, like the Time-of-Flight (ToF) sensors are widely explored in both medical and non medical fields with a capability of becoming part of computer aided diagnosis and surgery set-ups. Despite that, the 3D point clouds acquired using ToF sensors require advanced processing techniques and their analysis still poses a research challenge.

Soft body is one of physics engines built in computer vision software. It simulates closed surface models so that they look like soft tissue. Presented soft body technique is based on Runge-Kutta correct physics algorithm [4]. Custom edges of the model

J. Juszczyk (✉) · J. Czajkowska · B. Pycinski · E. Piętka
Faculty of Biomedical Engineering, Silesian University of Technology,
Zabrze, Poland
e-mail: jan.juszczyk@polsl.pl

are treated as a springs. The spring parameters, i.e. spring length, press and pulling force can be customized.

The aim of the study is to model the deformation of human body surface on the basis of data acquired by a ToF camera and the optical tracker. The skin surface region is imaged by the ToF camera during the experiment and the deformation is generated by a thin plastic stick. The acquired ToF data is synchronized with the trajectory signal received by the tracker. Then, the ToF point cloud is filtered, preprocessed, and analysed. The skin is approximated using Poisson surface reconstruction and the structure deformation is modelled by the soft body physics engine. The experiment is to evaluate the accuracy of the ToF data acquired during a clinical procedure that requires a skin surface deformation.

The paper is organized as follows. The next section introduces the experimental set-up used to evaluate the developed modelling technique and the applied methodology including: ToF data analysis, signals synchronization and dynamic model. The obtained results are given in Sect. 3, whereas the conclusions are stated in Sect. 4.

2 Materials and Methods

2.1 Experimental Set-Up

The carried out experiments aim at modelling of the human skin surface deformation under the pressure of a moving stick. The experimental set up consists of two data sources visualizing the object of interest: SwissRanger SR4000 ToF-camera¹ with the acquisition frequency of 10 Hz located in the distance of 1 m from the visualized object and the Polaris Spectra navigation system² with the acquisition frequency set to 25 Hz. The latter consists of a stereoscopic camera (i.e. the tracker) and the cross-shape marker rigidly attached to the stick. Tracking accuracy of the navigation system's markers is not worse than 0.25 mm according to the producer's note. Therefore, the location of the stick tip (recorded by the tracker) that measures the deformation depth is treated as ground truth.

For data processing two open source software are employed. PLUS toolkit allows the data to be merged [16]. Blender is a complete tool for creating and editing the 3D animation, games and movies, with build-in: game and render engines, movie editing and compositing algorithms, complete rigid body physics system and several physics engines. The software is widely used in various applications in current research and scientific simulations [8], also for medical application [17]. Digital models and phantoms of human body, and internal organs are developed for calculation of radiation risks and radiation protection in dosimetry [3, 15] or for surgical planning [12].

¹MESA Imaging 107 AG, Switzerland, <http://www.mesa-imaging.ch>.

²Northern Digital Inc., ON, Canada, <http://ndigital.com>.

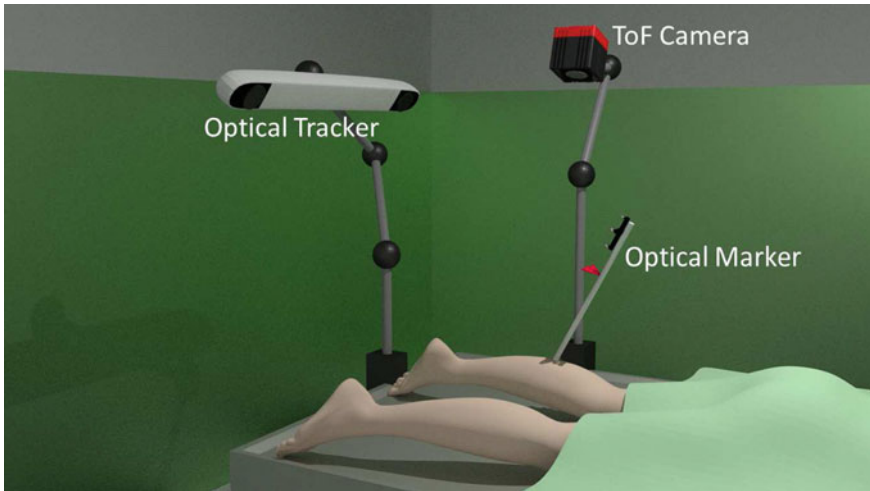


Fig. 1 Experimental set-up

The experimental set-up is shown in Fig. 1. The ToF-sensor records the point cloud of visualized object surface (i.e. the skin), whereas the optical positioning system records the accurate position of the stick.

2.2 ToF Data Analysis

The ToF camera measures the depth of a 3D scene. The imaging idea is based on multi-detector (charge coupled device—CCD) measurement of the optical signal generated by the device—infra-red light [14]. The ToF sensor module measures the distance by observing the phase shift of a modulation envelope of the light source. The scene is represented by the point cloud, where the Cartesian coordinates reflect the distance to the camera. The depth resolution of the camera is in the range of millimetres and does not require any computationally complex post-processing [6]. However, the ToF measurements still meet substantial challenges: low image and depth resolution, systematic or intensity-related distance error, depth inhomogeneity, motion artefacts, multiple scene reflections, blind zones in concave objects [14].

Preprocessing The ToF sensor produces a matrix data of depth, amplitude and a confidence value for every acquired point. The ToF data processing starts with the region of interest selection. Since the pressure of the moving stick influences the tissues in the neighbourhood only, the region to be analysed is limited. Thus, the automated detection of the stick end is performed on the amplitude image. To be sure that the stick touches the skin, the detection procedure starts with the middle image of the acquired set. The stick detection procedure uses the fuzzy c-means clustering and the Hough transform [2]. The region of interest is restricted to the circular area of

the radius depending on the modelled structure. Points that reflect the stick are found in the amplitude image and removed from the cloud, so that the data includes points corresponding to the actually deformed skin only.

To reduce the influence of artefacts and noise the ToF data is blurred. As a result every new point cloud is the mean of 5 consecutive clouds from the input data.

2.3 Data Synchronization

The experimental environment is stationary i.e. the position of the tracker, ToF camera and the phantom (except its surface) are invariant, and the stick is the only element that moves during the experiment. Since the mutual position of ToF camera and the tracker are constant, therefore there is no need to compute the spatial relations between their coordinate systems. However, since in our set-up the signals are recorded independently, the synchronisation between the tracker and ToF signals is required. The synchronisation allows the accuracy of the ToF data to be evaluated

The Algorithm of Data Synchronization Since the optical tracker calculates the position of the marker located on the stick, the position of the stick tip has to be found by the preceding calibration process [19]. It starts with the acquisition of a set of the stick positions obtained by its rotation the stick at various angles. Due to the fixed location of the stick tip, the resulting overdetermined system of linear equations yield the relation between stick marker and the tip [18].

Once the calibration is performed, the synchronisation of the ToF single point and the spatial trajectory of the stick tip precedes the skin press. It allows for comparison of the trajectory registered by the ToF and the optical tracker.

Data synchronisation is performed in four points. (1) First and last frames are chosen in both analysed signals, (2) Principal Component Analysis (PCA) dimensionality reduction technique [10] is employed for each signal to compute the main axis of movement, (3) the obtained signals are normalized to get mean value and variation equal to 0 and 1, respectively, (4) the time offset is computed by minimizing the root mean square errors between the signals.

2.4 Soft Body Modelling

In this section we present the generation of a model spread over the ToF data. It starts with the vertex normals calculation (Fig. 2) necessary for the mean weighted algorithm [9] to be employed. Then, the surface model is build by implementing the Poisson equation solution [13]. Due to the lack of the vertices distribution uniformity and unclosed surface in small curvature area, a mesh rebuilding step is included. It is based on a dual contouring method [11]. The created surface model is then passed to Blender soft body physics engine (Fig. 3).

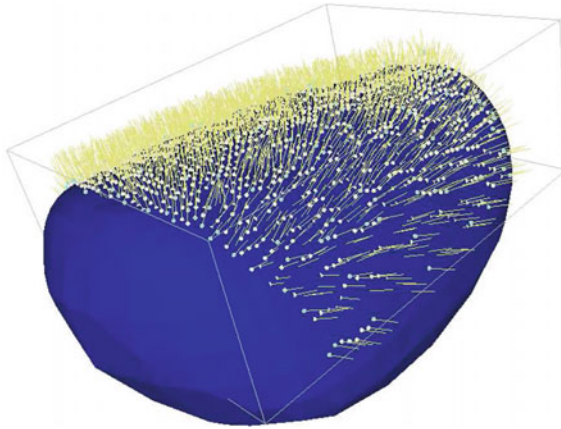


Fig. 2 The Poisson surface model (blue) and ToF points data with calculated normals (yellow)

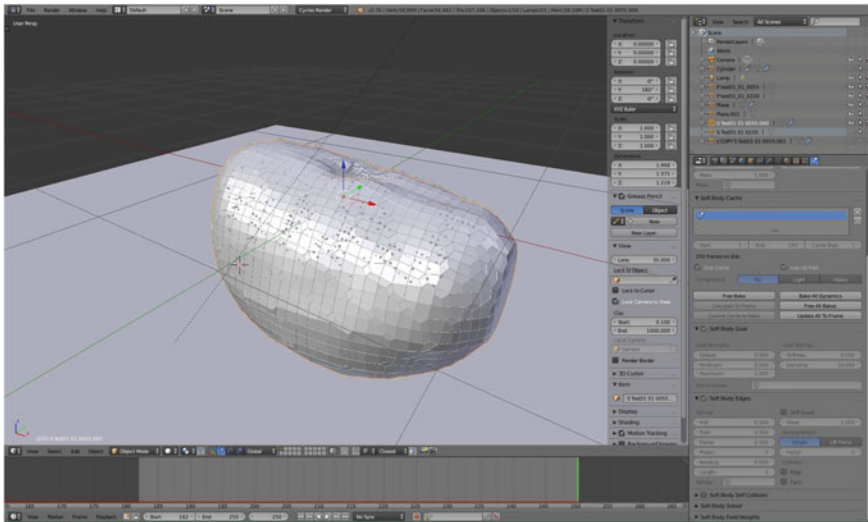


Fig. 3 Scene with Soft Body model in Blender software

3 Results

In this study, preliminary results of 3 experiments were presented. In each set-up, different positions of the stick caused various skin surface deformations.

The simulation starts with no press and continues until the maximal press. The data is acquired by the ToF camera as well as the optical tracker. The model was evaluated at both extreme positions and a middle one. The graphical representation of both extremes is shown in Fig. 4.

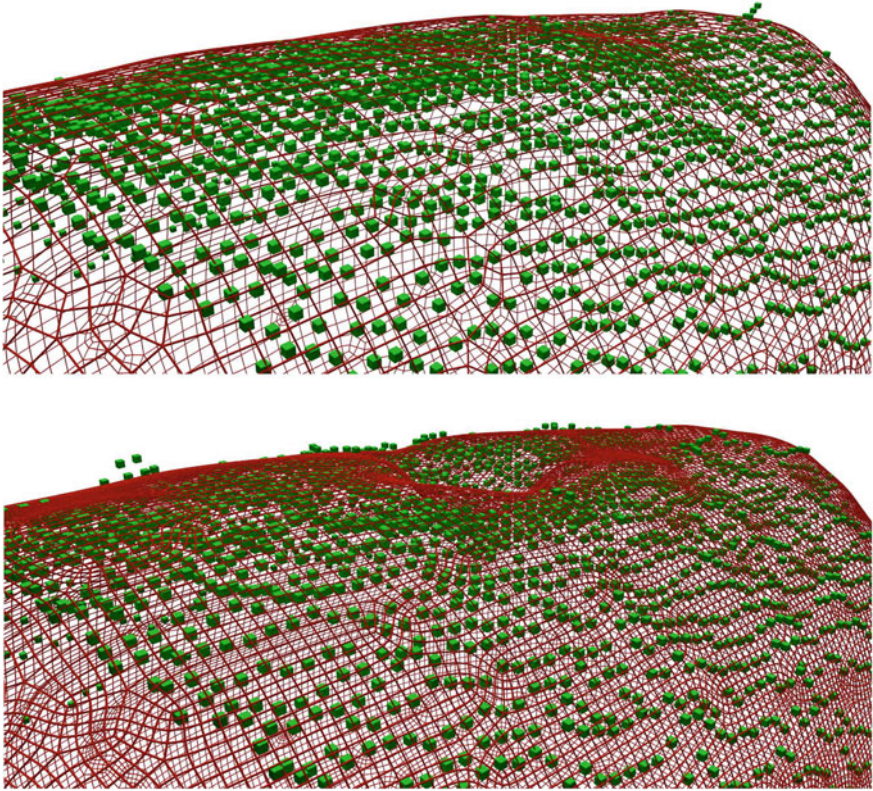


Fig. 4 Fusion of ToF-camera point cloud (*green dots*) and the model (*red grid*). *Top* without deformation, *bottom* at the moment of maximal deformation

For evaluation of the model accuracy, the directed Hausdorff distance H was employed [7]. It is defined as

$$H(X, Y) = \max_{x_i \in X} \min_{y_j \in Y} \|x_i - y_j\|, \quad (1)$$

where $\|\cdot\|$ is the Euclidean norm and $X = \{x_1, \dots, x_m\}$, $Y = \{y_1, \dots, y_n\}$ are two finite point sets. In this study X denotes ToF data, whereas Y is the set of vertices of the model.

In order to evaluate the model, the Soft Body Engine options: Soft Body Goal and Soft Body Edges were enabled, and parameters shown in Table 1 were adjusted experimentally.

To verify the usefulness of the mean filtering procedure in the pre-processing step, the directed Hausdorff distance was calculated for each pair in the set of every 5 consecutive point clouds in the acquired ToF set. The obtained mean value H was

Table 1 Parameters of Soft Body Engine used for the experiments

Parameter group	Parameter name	Value
Goal strengths	Default	0.90
Goal settings	Stiffness	0.50
	Damping	10.00
Springs	Pull/Push	0.50
	Bending	0.00

equal to 1.2 mm, whereas the mean H value between the first and the fifth cloud in the analysed set was equal to 1.5 mm.

In order to detect the largest skin deformation in ToF-recording, the ToF and optical tracker synchronization step was applied. The accuracy of $H = 0.15$ mm was estimated by calculating Hausdorff distances between the trajectory of the marker position found in ToF data and the trajectory along the main axis (as described in Sect. 2.3) registered by the optical tracker (Fig. 5).

The final evaluation was performed at this axial slice, where the depth of deformation reached the maximum. H values were calculated in three positions: when the stick touched the skin surface, in the middle position and when reached the strongest pressure (Fig. 6).

In the deformation area, the Hausdorff distances between the ToF images and the model did not exceed 3.0 mm (Table 2).

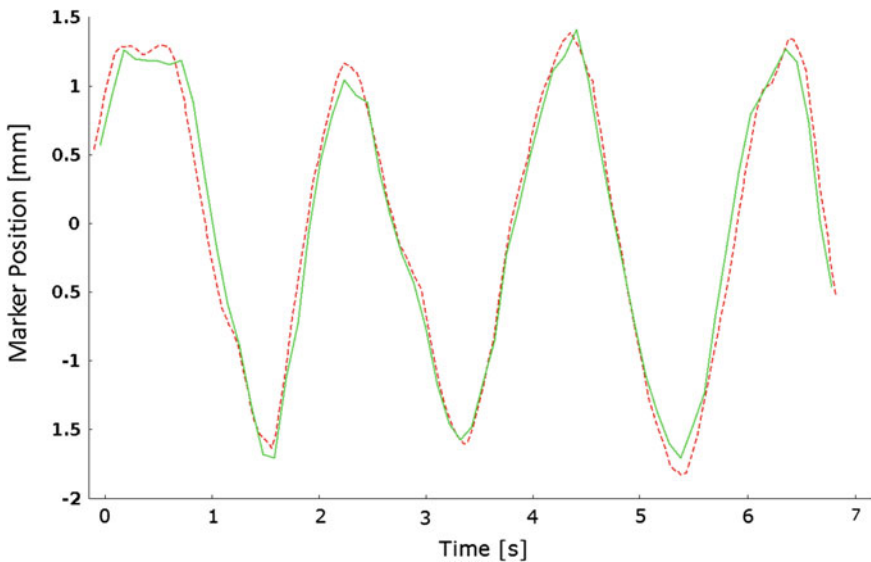


Fig. 5 ToF (green solid line) and optical tracker (red dashed line) data synchronization

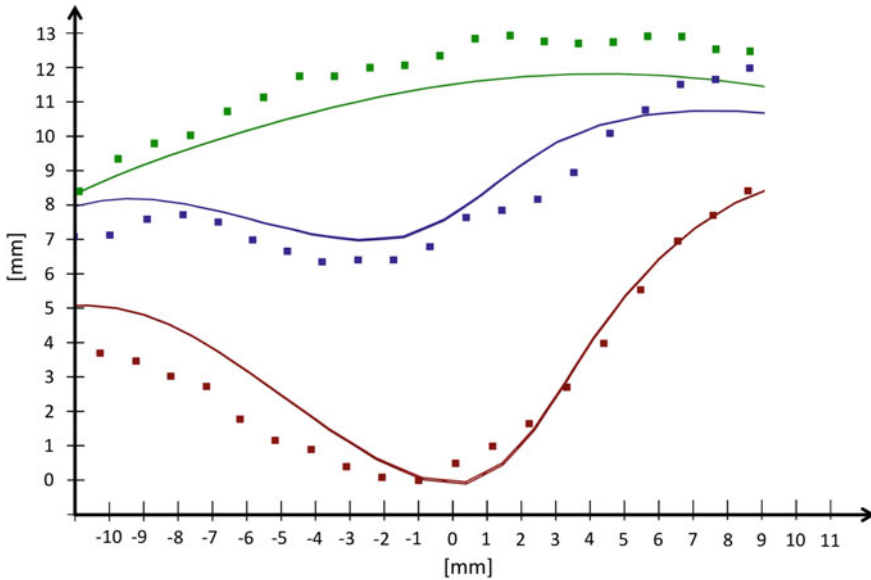


Fig. 6 Comparison of the ToF acquisition (*dots*) and the model (*solid line*) of a surface of single axial slice registered in three pressure positions: no pressure value—*green*; 50% of maximal pressure value—*blue*; maximal pressure value—*red*

Table 2 H values between ToF data and the developed model at three moments of three deformation experiments (mm)

Experiment no.	Begin	Median	Maximum
1	1.7	1.6	2.2
2	1.2	3.0	2.3
3	2.0	2.4	2.2

4 Conclusion

The paper aims at modelling human skin surface and subcutaneous tissue deformation under the influence of external pressure. The main idea of the methodology is to apply the Poisson surface reconstruction and soft body physics engine in order to model the point cloud registered by the ToF-sensor. The developed method has been evaluated using the image data of the skin surface deformed by the plastic stick. Time of computation the model deformation does not exceed one second per frame.

The promising results prove the robustness of the presented approach and encourage us to open up a wide study on the internal tissue dynamic modelling by using the physics engines research. The paper is a preliminary study to be implemented on other soft tissue surfaces including the abdominal skin surface modelling during a trocar puncture at the beginning of image navigation aided laparoscopical surgery.

Acknowledgments This research was supported by the Polish National Science Center (NCN) grant No. UMO-2012/05/B/ST7/02136.

References

1. Bugdol, M., Czajkowska, J., Pietka, E.: A novel model-based approach to left ventricle segmentation. In: Murray, A. (eds.): *Computing in Cardiology (CINC)*, vol 39. *Computers in Cardiology Series*, pp. 561–564 (2012)
2. Czajkowska, J., Pyciński, B., Piętka, E.: HoG feature based detection of tissue deformations in ultrasound data. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6326–6329 (2015)
3. de Cassola, V., Melo Lima, V., Kramer, R., Khoury, H.: Fash and mash: female and male adult human phantoms based on polygon mesh surfaces: I. development of the anatomy. *Phys. Med. Biol.* **55**(1), 133 (2009)
4. Fonseca, T.C., Lebacqza, A.L., Mihailescu, L.C., Vanhaverea, F., Bogaerts, R.: Development of a 3D human body library based on polygonal mesh surface for whole body counter set-up calibration. *Progress in Nuclear Science and Technology (PNST)*. Nara Prefecture: Atomic Energy Society of Japan (2014)
5. Geng, J.: Structured-light 3d surface imaging: a tutorial. *Adv. Opt. Photon.* **3**(2), 128–160 (2011)
6. Gokturk, S., Yalcin, H., Bamji, C.: A time-of-flight depth sensor—system description, issues and solutions. In: *Conference on Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04*, pp. 35–35 (2004)
7. Huttenlocher, D., Klanderman, G., Rucklidge, W.: Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863 (1993)
8. Izadi, E., Bezuijen, A.: Simulation of granular soil behaviour using the bullet physics library. In: *International Symposium on Geomechanics from Micro to Macro (TC 105 ISSMGE)*, vol. 2, pp. 1565–1570 (2015)
9. Jin, S., Lewis, R.R., West, D.: A comparison of algorithms for vertex normal computation. *Vis. Comput.* **21**(1–2), 71–82 (2005)
10. Jolliffe, I.T.: *Principal Component Analysis*, 2nd edn. Springer, New York (2002)
11. Ju, T., Losasso, F., Schaefer, S., Warren, J.: Dual contouring of hermite data. *ACM Trans. Graph. (TOG)* **21**(3), 339–346 (2002)
12. Juszczyk, J., Pyciński, B., Piętka, E.: Patient specific phantom in bimodal image navigation system. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2908–2911 (2015)
13. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, vol. 7 (2006)
14. Kolb, A., Barth, E., Koch, R., Larsen, R.: Time-of-flight cameras in computer graphics. *Comput. Graph. Forum* **29**(1), 141–159 (2010)
15. Kramer, R., Khoury, H., Cassola, V., de Melo Lima, V.: Preview of a series of adult human phantoms for radiation protection dosimetry. In: *World Congress on Medical Physics and Biomedical Engineering*, pp. 120–123, 7–12 Sept 2009. Springer, Munich (2009)
16. Lasso, A., Heffter, T., Rankin, A., Pinter, C., Ungi, T., Fichtinger, G.: PLUS: open-source toolkit for ultrasound-guided intervention systems. *IEEE Trans. Biomed. Eng.* **61**(10), 2527–2537 (2014)
17. Milusheva, S., Karastanev, S., Toshev, Y.: Open Source Information Technologies Approach for Modeling of Ankle-Foot Orthosis. *Institute of Information Theories and Applications FOI ITHEA* (2007)
18. Onprasert, W., Ongwattanakul, S., Suthakorn, J.: Implementation on a new tool tip calibration method for biomedical applications. *Recent Adv. Comput. Sci. Inf. Eng.* **129**, 385–392 (2012)

19. Pyciński, B., Juszczuk, J., Bożek, P., Ciekalski, J., Dzielicki, J., Piętka, E.: Image navigation in minimally invasive surgery. In: *Information Technologies in Biomedicine*, vol. 4. *Advances in Intelligent Systems and Computing*, vol. 284, pp. 25–34 (2014)
20. Swiatek-Najwer, E., Bedzinski, R., Krowicki, P., Krysztoforski, K., Keppler, P., Kozak, J.: Improving surgical precision-application of navigation system in orthopedic surgery. *Acta Bioeng. Biomech.* **10**(4), 55–62 (2008)

Evaluation of Radiofrequency Field (2.4 GHz) Effects to Multilayer Structure of Human Skin and Pacemaker

Zuzana Psenakova, Maros Smondrk, Grazia Lo Sciuto
and Mariana Benova

Abstract The proposed paper deals with simulation and analysis of the radiofrequency field (2.4 GHz) inside the human tissue using a multi-layered skin and pacemaker model. The research was carried out using electromagnetic modelling based on the Finite Integration method. The simulations were performed in terms of computing electric field distributions calculated for four different hypodermis layer thickness. The focus was on near-field exposure modelled by dipole antenna. The results have shown that higher hypodermis layer causes higher absorption of electromagnetic field which leads to lower values of electric field strength on the pacemaker casing.

Keywords Finite integration method · Electromagnetic field · Radiofrequency radiation · Pacemaker · Skin tissue

1 Introduction

The increased use of wireless technologies in last decade leads to increased level of electromagnetic radiation to which human are exposed. These technologies as a source of high-frequency electromagnetic fields involve devices such as mobile phones, digital cordless phones, Bluetooth devices, Wi-Fi equipment and so on. Since the potential physiological effects of radiofrequency radiation and its thermal and non-thermal effects are still under scientific investigation, the specific absorption rate (SAR) or the induced current density in human body is evaluated for health risk assessment. In general, it is very difficult to measure internal electric field intensity or temperature elevation in the human body using non-invasive methods. However, the application of boundary condition method to the Infrared Thermography can be

Z. Psenakova (✉) · M. Smondrk · M. Benova
Faculty of Electrical Engineering, Department of Electromagnetic
and Biomedical Engineering, University of Zilina, Zilina, Slovakia
e-mail: zuzana.psenakova@fel.uniza.sk

G.L. Sciuto
Department of Industrial Engineering, University of Catania, Catania, Italy

used in order to find internal temperature distribution, the common way to do this is to use modelling tools such as an electromagnetic simulations software.

Previous research conducted over last decade has demonstrated the complexity of interaction of electromagnetic energy with a human tissue. The human tissue is usually represented as a multi-layered structure [16, 17]. The interaction between incident electromagnetic wave and human tissue depends on the nature of the tissue by means of its electromagnetic properties which vary mainly with change of the radiation frequency [7, 9]. Dosimetric studies quantify these interactions and define standards and exposure limits in terms of the SAR (a measure of the rate at which energy is absorbed by the human tissue while exposed to a radiofrequency electromagnetic radiation). SAR is defined as follows:

$$SAR = \frac{\sigma \cdot |E|^2}{\rho}, \quad (1)$$

where $|E|$ —maximal value of intensity of the electric field ($V m^{-1}$) in the tissue, σ —electrical conductivity ($S m^{-1}$) of the tissue, ρ —mass density of the tissue ($kg m^{-3}$), [8, 18]. To estimate SAR value in human tissue at certain depth, the intensity of electric field or the electric field strength must be known. Distribution of the electric field strength E is defined as follows:

$$\nabla \times \mu_r^{-1} (\nabla \times E) - k_0^2 \left(\varepsilon_r - \frac{j\sigma}{\omega \varepsilon_0} \right) E = 0, \quad (2)$$

where k_0 —is wave vector in free space (m^{-1}), ∇ —rotation vector operator, ε_r —relative permittivity, ε_0 —vacuum permittivity ($F m^{-1}$), μ_r —relative permeability, σ —electrical conductivity ($S m^{-1}$) and ω —angular wave frequency ($rad s^{-1}$), [8, 18]. Knowing these quantities and physical parameters of exposed tissue, it is possible to evaluate the specific absorption rate (SAR) of the particular tissue at certain depth by using the Eqs. (1, 2). However, the distribution of electric field strength inside multi-layered structure is more complex by means of electromagnetic wave reflections on each layer and their mutual interaction. Moreover previous research have shown that metallic implants may change the distribution of electric field strength inside the tissue [1]. Metallic implants may thus modify absorption of electromagnetic energy around them by means of scattering of the incident electromagnetic wave. In case when the metallic implants such as a pacemaker lies in the surface layers of the human body, the pacemaker casing may lead to enhancement of the electric field strength and SAR inside the body.

In this context, the aim of this study was to model, simulate and analyse the electric field distribution inside the multi-layered structure (comprising a skin and a pacemaker model) exposed to 2.4 GHz radiofrequency radiation. The study was performed in terms of electric field distributions computed at four different hypodermis layer thickness.

2 Skin Tissue and Metallic Implant

The human skin is composed of three basic layers: epidermis, dermis and hypodermis (Fig. 1). Hypodermis, also known as subcutaneous tissue or fat, lies immediately below the dermis. Subcutaneous tissue layer consist of many cell types like fibroblasts, fat cells (adipocytes) and leukocytes.

The subcutaneous tissue fulfill mainly the function of mechanical protection against pressure and friction, energy reserve and thermal regulation. The thickness of hypodermis varies considerably in respect of body area, nutritional status, age and gender of the individual. In thin areas of the skin like eyelids, genitals, and nipples, it is nearly absent. With ageing drastic changes take place in the hypodermis. It begins to thin out and disintegrate causing deep wrinkles, folds and sagging of the facial skin. The adult’s hormones have effect on the pattern of hypodermis fat distribution function. For example in men fat tends to accumulate over the abdomen, torso and shoulders and in women fat accumulates usually on the breasts, waist, hips, buttocks and thighs. In weight loss there is usually changes in the amount of lipid content of the adipocytes and not in their number. The important value in this case is the Body Mass Index (BMI). BMI is a calculation, based on body height and weight as to the amount of fat on your body. Depending on the result, person can be classified as an underweight, normal, overweight, and obese [10].

In this study, we have focused on the thickness of the hypodermis which considerable changes with the aforementioned BMI classification. The thickness values of hypodermis for different BMI classes are reported in the Table 1.

A pacemaker is the implantable therapeutic device providing cardiac pacing by means of long-term intrinsic cardiac rhythm support or replacement. It is surgically

Fig. 1 The main layers of the human skin

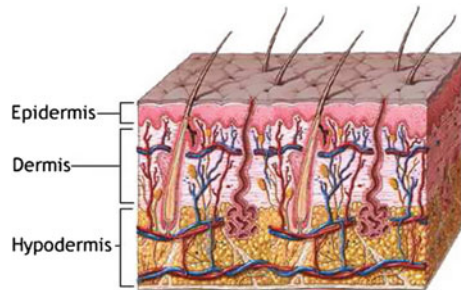
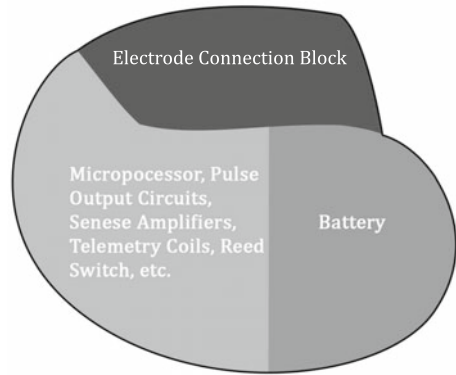


Table 1 Hypodermis layer in the chest near shoulder area of people with different BMI [10]

BMI	Man hypodermis (mm)	Woman hypodermis (mm)
Underweight	<3	<7
Normal	3–5	7–10
Overweight	6–16	11–20
Obese	>16	>20

Fig. 2 Simple pacemaker diagram [14]



inserted in the patient’s body directly under the skin tissue in the chest near left shoulder area where an incision is made below the collar bone creating a small pocket. Pacemaker is an electronic device comprising of a pulse generator, battery, internal electrical circuits, electrical coils for telemetry and so on (Fig. 2).

Since pacemaker is an active electronic device which itself is a source of electromagnetic field, in this study we have considered a pacemaker as a passive metallic implant. The pacemaker casing is usually made from a biocompatible titanium alloys such as Ti-Zr and Ti-Nb [5, 14, 15].

3 Exposure Limits

The International Commission on Non-Ionizing Radiation Protection (ICNIRP) sets exposure limits for EMF fields. Our interest is focused on E-field strength for frequency between 2 and 300 GHz. In tissue, SAR is proportional to the square of the internal electric field strength. Average SAR and SAR distribution can be computed or estimated from laboratory measurements [2, 3]. Values of SAR depend on the following factors (Table 2):

- the incident field parameters, i.e., the frequency, intensity, polarization, and source-object configuration (near- or far-field);

Table 2 Basic restrictions for E-field strength (2–300 GHz) and SAR values (10 MHz–10 GHz) [11]

Exposure characteristics	E-field strength ($V m^{-1}$)	Whole-body average SAR ($W kg^{-1}$)	Localized SAR (head and trunk) ($W kg^{-1}$)	Localized SAR (limbs) ($W kg^{-1}$)
Occupational	137	0.4	10	20
General public	61	0.08	2	4

- the characteristics of the exposed body, i.e., its size and internal and external geometry, and the dielectric properties of the various tissues;
- ground effects and reflector effects of other objects in the field near the exposed body [11].

4 Numerical Modelling of Biological Tissue

This study is based on electromagnetic modelling using the Finite Integration Technique (FIT) implemented within a commercially available simulation software the CST Microwave Studio. The time-domain solver based on the FIT was applied on hexahedral mesh, on account of the software recommendation for dipole antennas application [6].

A multi-layer tissue model comprising skin tissue and pacemaker casing was designed (Fig. 3). The multi-layered structures and their thickness were derived in terms of previous chapter devoted to skin tissue. Since the precise choice of implant pocket location and depth is based upon physician preference and cannot be controlled, we consider its location directly under the skin. However, other authors have considered pacemaker location within tissue model comprising a skin, fat and muscle layer [4, 12, 16]. We assume its most superficial location, thus the proposed model consists of epidermis, dermis, subcutaneous tissue and metallic implant. The dielectric properties of skin layers were taken from database [9]. The material considered for the implantable pacemaker casing was a titanium with a thickness of the

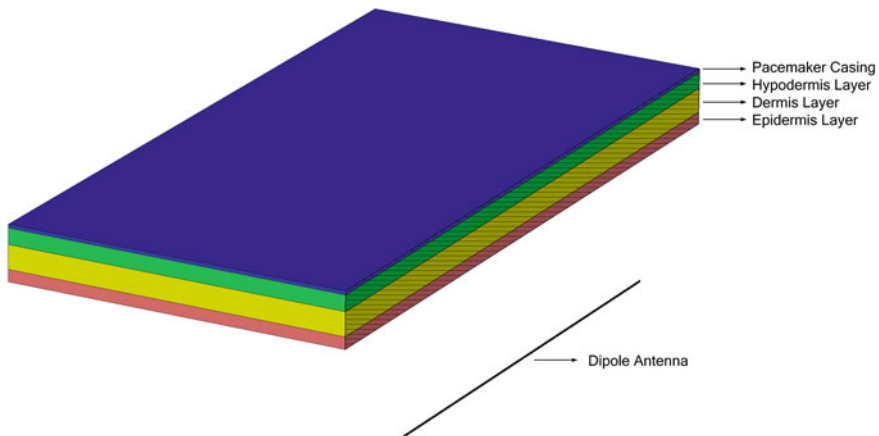


Fig. 3 The geometrical layout of a used multilayer structure model and dipole antenna. The perspective view with vertical cutting plane crossing the longitudinal axis of the dipole antenna was used to emphasize the mutual distance between the antenna and dermis. It was set up to 15 mm. *Black* cylinder structure is a dipole antenna, *red block* is a epidermis layer, *yellow block* is a dermis layer, *green block* is a hypodermis layer and *blue block* is a pacemaker layer

Table 3 Thickness and dielectric properties of the multilayer model (frequency of 2.4 GHz) [9, 13]

Layer	Thickness (mm)	ε_r	$\sigma(\text{S m}^{-1})$
Epidermis	1.5	38.1	1.44
Dermis	3	25.2	5.76
Hypodermis	2/5/10/18	30.5	3.6
Pacemaker casing	0.45	50	0.267

housing wall equal to 0.45 mm [4]. The thickness and dielectric properties of each individual layer are summarised in Table 3.

The layers of epidermis, dermis and metallic implant maintain constant thickness. We varied the thickness of the hypodermis in terms of BMI classes: 2 mm for underweight, 5 mm for normal, 10 mm for overweight, and 18 mm for obese. Tuned-dipole antenna was used as a source of near-field electromagnetic exposure. The antenna was modelled as a discrete port with a reference impedance of 50Ω and Gaussian shaped excitation signal. The multilayer structure was irradiated by 2.4 GHz radiofrequency wave with power of 0.5 W. The simulation domain was extended to include the air—multilayer interface, so the antenna was placed static 15 mm from the first skin layer.

5 Results and Discussion

The simulation was performed for four different hypodermis layer thickness according to the four basic BMI classes. The calculated distributions of electric field strength within multilayer structure for all cases are shown in the contour plots within cutting plane crossing the longitudinal axis of the dipole antenna (Fig. 5). The cutting plane crossing the longitudinal axis of the dipole antenna was chosen, on account of the fact that there is localised the maximum power radiated by antenna. Similarly, the absolute value of electric field strength on curves crossing the tissue layers in the exact centre of the computation space (perpendicular to the longitudinal axis of the dipole antenna) for all BMI classes are shown in Fig. 4.

The highest electric field strength at the hypodermis-pacemaker interface was observed for hypodermis thickness of underweight man, (Figs. 5A, 4). The absolute value of electric field strength at the hypodermis-pacemaker interface was 59.5 V m^{-1} . For the hypodermis layer thickness of the normal man (Figs. 5B, 4), the absolute value of electric field strength was lower than underweight man condition, equals to 40.3 V m^{-1} . The lower electric field strength was observed for case of overweight man, (Figs. 5C, 4). The absolute value of electric field strength at the hypodermis-pacemaker interface was 21.7 V m^{-1} .

The lowest distribution of electric field strength was observed for hypodermis thickness of obese man, (Figs. 5D, 4). The absolute value of electric field strength at the hypodermis-pacemaker interface for obese case was 8.1 V m^{-1} . Due to the

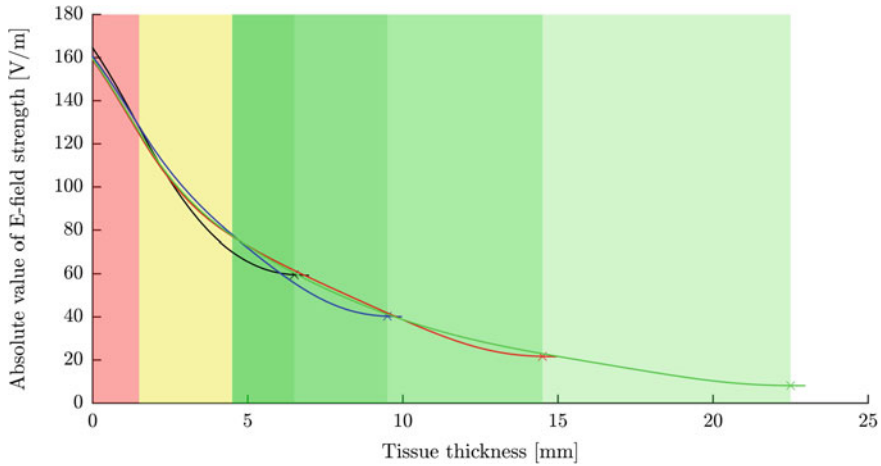


Fig. 4 The distribution of electric field strength on curves crossing the tissue layers in the exact centre of the computation space (perpendicular to the longitudinal axis of the dipole antenna) for all BMI classes: underweight—black line, normal—blue line, overweight—red line, obese—green line. The cross marks denote the hypodermis-pacemaker interface in same manner: underweight—black cross, normal—blue cross, overweight—red cross, obese—green cross. Additionally, the coloured rectangles indicate the tissue layer of proposed model as follows: red—epidermis, yellow—dermis, green—hypodermis which thickness varies with respect to the BMI classes

symmetry of the dipole antenna leading to the symmetry of electric field distribution shown in Fig. 5 and anisotropic material properties of designed model, the maximal electric field strength within tissue model lies mainly in the centre of computation space. Thus maximal values of electric field strength for all BMI classes are localised on curves shown in Fig. 4. The electric field strength for all BMI classes at the air-epidermis interface slightly ranges from 159 to 165 $V\ m^{-1}$ regardless the constant thickness of air and epidermis layer, on account of the fact that the adaptive mesh refinement was used for each model. The distribution of electric field strength within tissue model have revealed the declining tendency as the result of lossy tissue environment which attenuate electromagnetic field. Therefore, the electric field strength at hypodermis-pacemaker interface have shown declining tendency while increasing the hypodermis layer thickness. The local increase of electric field strength in tissue layers above to the hypodermis-pacemaker interface as a result of electromagnetic wave reflections at this interface were not observed, mainly due to the low electric conductivity of pacemaker casing material. Additionally, the calculated values of electric field strength within multi-layer model for all BMI classes vary within restriction limits for the occupational as well for general public exposure.

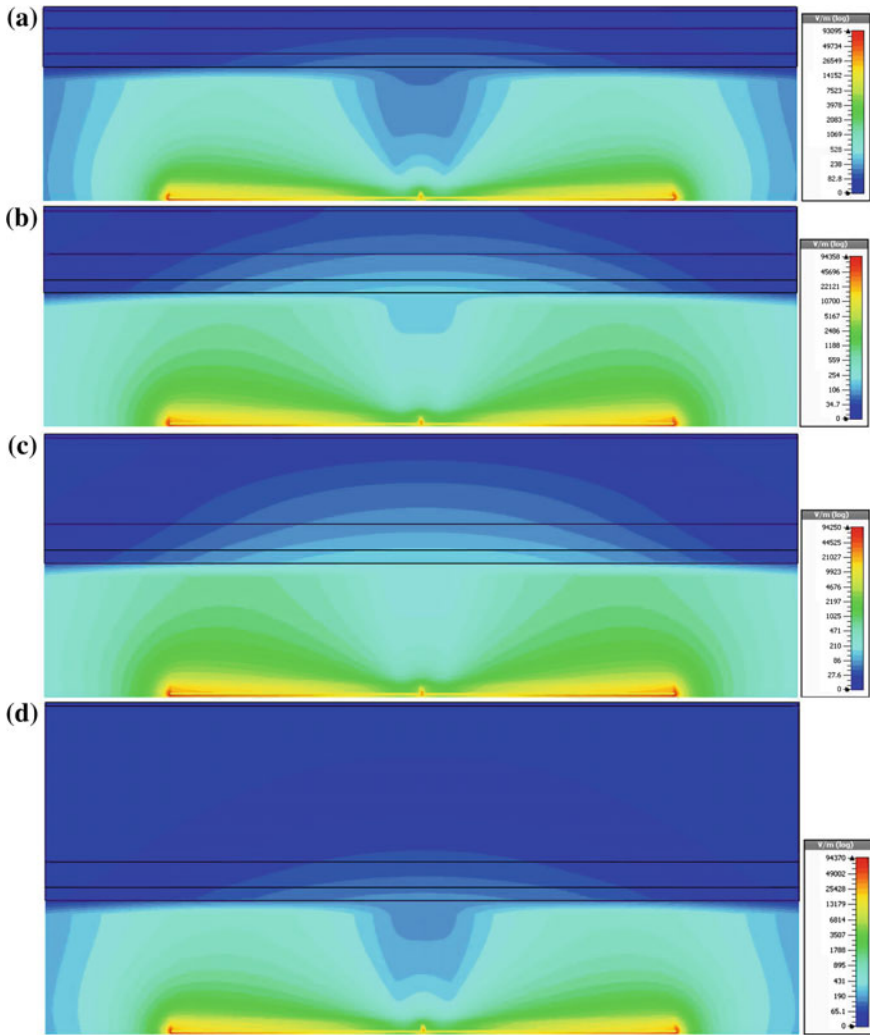


Fig. 5 The distribution of electric field strength in the multilayer model within cutting plane crossing the longitudinal axis of the dipole antenna and the centre of the computation space for four BMI classes: underweight (a), normal (b), overweight (c), and obese (d) man

6 Conclusion

The simulation of electric field strength inside the multi-layered model of human skin and pacemaker casing have shown that the absolute value of electric field strength vary with respect to the hypodermis layer thickness. Thicker hypodermis layer or subcutaneous fat causes higher absorption of propagated electromagnetic field which leads to lower values of electric field strength within tissue model as well as on the

pacemaker casing. It is important to know own BMI or classes such as underweight, normal, overweight, and obese, because having too much body fat are can cause many health related issues. On contrary, simulation results have confirmed the fact that more subcutaneous fat reduce electric field strength to which metallic implant could be exposed. Thus could protect the pacemaker before possible negative effects of external electromagnetic field.

Acknowledgments This work has been supported from project APVV-14-0519 “Smart textiles and clothing for mobile monitoring of human vital functions”.

References

1. Anderson, V., McIntosh, R.: Guidelines for the RF Exposure Assessment of Metallic Implants. MCL (2006)
2. Association for the Advancement of Medical Instrumentation: Active implantable medical devices—Electromagnetic compatibility—EMC test protocols for implantable cardiac pacemakers and implantable cardioverter defibrillators. Technical Report ANSI/AAMI PC69:2007 (2007)
3. Calcagnini, G., Censi, F., Bartolini, P.: Electromagnetic immunity of medical devices: the European regulatory framework. *Ann. Ist. Super. Sanita* **43**(3) (2007)
4. Campi, T., Cruciani, S., Santis, V.D., Feliziani, M.: EMF safety and thermal aspects in a pacemaker equipped with a wireless power transfer system working at low frequency. *IEEE Trans. Microwave Theory Techniq.* (2016)
5. CENELEC European Committee for Electrotechnical Standardization: Procedure for the assessment of the exposure to electromagnetic fields of workers bearing active implantable medical devices: Specific assessment for workers with cardiac pacemakers. Technical Report BS EN 50527-2-1:2011 (2011)
6. CST Microwave Studio: Workflow and Solver Overview (2015)
7. Gabriel, S., Lau, R.W., Gabriel, C.: The dielectric properties of biological tissues: II. measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med. Biol.* **41**(11) (1996)
8. Golio, M.: The RF and Microwave Handbook, 2nd edn, vl. 3 Set. Electrical Engineering Handbook. CRC Press (2008)
9. Hasgall, P., Gennaro, F.D., Baumgartner, C., Neufeld, E., Gosselin, M., Payne, D., Klingensböck, A., Kuster, N.: IT’IS Database for thermal and electromagnetic parameters of biological tissues, V3.0. IT’IS Foundation (2015)
10. Holcombe, S., Wang, S.: Subcutaneous fat distribution in the human torso (2014)
11. International Commission on Non-Ionizing Radiation Protection: Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz). *Health Phys.* **74**(4) (1998)
12. Li, P., Zhang, L., Liu, F., Amely-Velez, J.: Optimized matching of an implantable medical device antenna in different tissue medium using load pull measurements. In: Microwave Measurement Conference, 2015 86th ARFTG (2015)
13. Pavselj, N., Miklavcic, D.: Numerical models of skin electropermeabilization taking into account conductivity changes and the presence of local transport regions. *IEEE Trans. Plasma Sci.* **36**(4) (2008)
14. Psenakova, Z., Hudecova, J.: Influence of electromagnetic fields by electronic implants in medicine. *Elektronika ir Elektrotechnika* **95**(7) (2015)
15. Tandogan, I., Temizhan, A., Yetkin, E., Guray, Y., Ileri, M., Duru, E., Sasmaz, A.: The effects of mobile phones on pacemaker function. *Int. J. Cardiol.* **103**(1) (2005)

16. Vidal, N., López-Villegas, J.M.: Changes in electromagnetic field absorption in the presence of subcutaneous implanted devices: minimizing increases in absorption. *IEEE Trans. Electromagn. Compat.* **52**(3) (2010)
17. Varotto, G., Staderini, E.M.: A 2D simple attenuation model for EM waves in human tissues: comparison with a FDTD 3D simulator for UWB medical radar. In: 2008 IEEE International Conference on Ultra-Wideband, Institute of Electrical and Electronics Engineers (IEEE) (2008)
18. Walendziuk, W.: Parallel computation of the SAR distribution in a 3D human head model. *Proc. SPIE* **6937** (2007)

Fluid Flow Analysis Using Finite Element Method, Determining the Effects of the Implantable Mechanical Heart Valves on Aortic Blood Flow

Przemyslaw Kurtyka, Witold Walke and Marcin Kaczmarek

Abstract The artificial heart valves have been already implanted for over 50 years, but still the perfect valve has not been designed yet. Work is focused on improving the parameters of the flow through the valve, but also the reduction of functional disorders of the cellular components of blood and perigraft tissue. The authors have attempted to analyze the blood flow through the two types of valves—bileaflet based on the St. Jude Medical On-X model and with tilting disc based on Moll's construction, which are currently used in treating heart failure. The flow analysis was performed using the FSI (Fluid Structure Interaction) module, which allows the determination of the characteristics of blood pressure and velocity at different opening angles of the valve. Information obtained during such analysis may be the basis to identify alternative biomaterials for mechanical heart valves, but also allows the observation of the phenomena of the blood flow.

Keywords FSI analysis · ON-X heart valve · MOLL heart valve · Blood flow velocity · Blood flow pressure

1 Introduction

Nowadays heart is more often the main cause of death in group of young people. The reason for this are not only the external factors such as diet, lack of physical activity or drugs, but also congenital heart defects resulting from prenatal disorders. Heart valves, which control the blood flow, are also susceptible to many dysfunctions. The occurrence of infections have the adverse effect to the mechanical properties of the heart valves [4–6]. However the rheumatic fever, endocarditis and degenerative diseases of connective tissue may also adversely affect the performance of the valve.

P. Kurtyka (✉) · W. Walke · M. Kaczmarek
Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
e-mail: przemyslaw.kurtyka@polsl.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_23

Currently, there are several methods used to counter the effects of valvular insufficiency. The first attempt of treatment is the drug therapy, which usually reduces only the effects and symptoms of the valve malfunction. The performance of a surgical procedure is more invasive method, but it allows the reparation of the natural valve or its replacement with a new one. The limited availability of homografts forced scientists to produce valvular replacement [9–12].

The artificial heart valves have been already implanted for over 50 years, but still the perfect valve has not been designed yet. The work is focused on improving the parameters of the flow through the valve, but also the reduction of functional disorders of the cellular components of blood and perigraft tissue. Scientists are looking for the material which is marked by improved resistance to mechanical and structural wear with simultaneous minimizing the risk of blood clots.

Nowadays multiple types of heart valves are used. They differ in the mechanism, design and the executive material. Currently ongoing work is focused on a suitable selection of the biomaterial, the modification method of the surface layer to improve hemocompatibility, but also a constructional solution providing and improving stability and effectiveness. One of the problems associated with the construction of mechanical heart valves is the method of fixing the disk in the valve housing. This method has a direct effect on proper blood flow through the valve [17, 18, 23]. The main problems that may occur regarding to the construction of the valve are hemolysis and the risk of embolic complications. It is associated primarily with the type of flow that is formed by an alternating over- and underpressure in the perigraft cavities and the opening angle of the valve during systole in the range of 60–90°. Use of mechanical heart valves is also associated with the risk of stress corrosion and the need for anticoagulant therapy [24–29]. However the risk of mechanical damage is much lower as compared to biological valves. Currently, the bileaflet and the tilting disc valves are the most widely used. In order to determine the suitability of the specific construction of the heart valve it is necessary to conduct the comparative analysis, which simulate the actual conditions in a biological system. For this purpose, numerical methods are used to simulate the phenomena occurring during implantation and implant exploitation. One of the methods to optimize the geometrical features and proper selection of the mechanical properties of implants is FEA (Finite Element Analysis). These analyzes are used regardless to the type of the implant and the place of implantation [1, 2, 14–16, 19–21, 30]. In addition to basic information obtained during FEA, especially for the implants localized in blood vessels, FSI analyzes are conducted. The obtained results allow the phenomena to be taken into account in relation to the blood flow through the geometry of the implant, which has a positive impact on improving its quality. Thus, the authors analyzed the influence of the specific constructions of the heart valve on blood flow and its impact on the components of the valve.

2 Methodology

2.1 Numerical Model of the Vessel—Valve System

The analysis covered two types of heart valves, bileaflet based on the St. Jude Medical On-X model and with tilting disc based on Moll’s construction, which are the most commonly used to treat heart failure (Fig. 1).

The 3D models of both heart valves were prepared in Solid Edge ST7 software basing on documentation provided by the Foundation for Cardiac Surgery Development in Zabrze (Fig. 2).

The analysis was performed for three characteristic opening angles of the valve i.e.: open, half-open, closed. Depending on the type of the valve, the opening angles were adopted based on the orientation of the discs in relation to the valve housing:

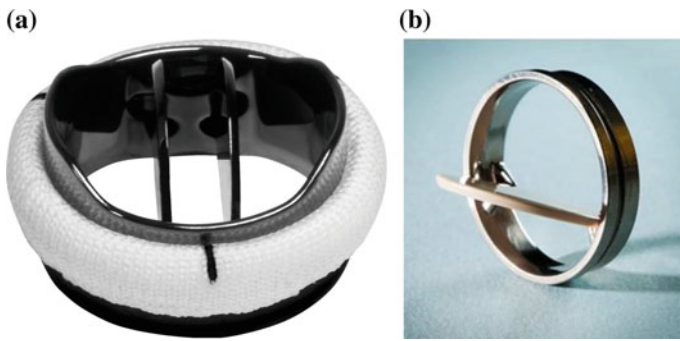


Fig. 1 The analyzed types of heart valves, **a** bileaflet—St. Jude Medical On-X, **b** tilt-disc valve—Moll type

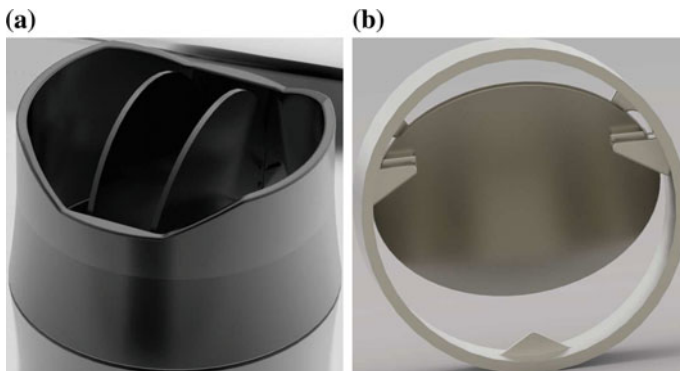


Fig. 2 The heart valve models designed in Solid Edge ST7 software, **a** bileaflet—St. Jude Medical On-X, **b** tilt-disc valve—Moll type

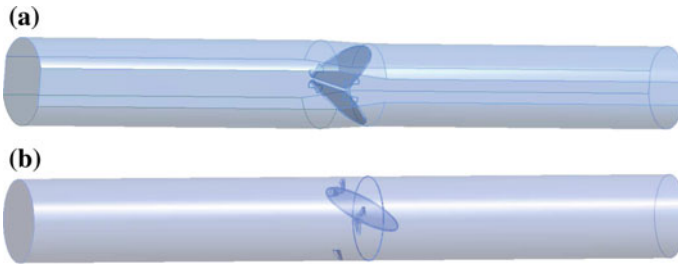


Fig. 3 The geometrical form of the fluid flow domain for different types of heart valves: **a** bileaflet (position 0%), **b** tilting disc (position 100%)

- for bileaflet heart valve, i.e.: open angle— 90° (position 100%), half-open angle— 60° (position 50%), closed angle— 40° (position 0%): Fig. 2a;
- for tilting disc heart valve, i.e.: open angle— 71° (position 100%), half-open angle— 40° (position 50%), closed angle— 40° (position 0%): Fig. 2b;

In a further step the flow domain was created. For this purpose the tubular geometry was built in the inlet and outlet of the valve. The wall thickness has been set as $g = 1$ mm for the length of the tube equal $l = 100$ mm (Fig. 3). During the simulation, models were simplified to overcome the risk of error and to obtain close approximation to the actual state. For this purpose sewing cuffs, which enable attaching to the heart tissue, were omitted.

2.2 Computational Model of the Vessel—Valve System

In the next stage of work, discretization was performed for all models, by dividing them for the finite elements of minimal size 0.11 mm (Fig. 4).

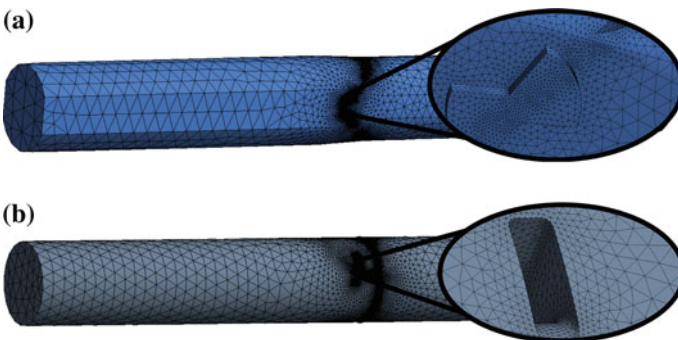


Fig. 4 Fluid domain after discretization for: **a** bileaflet mechanical valve, **b** tilting disc mechanical valve

Subsequently three surface groups were created: inlet and outlet indicating the direction of the blood flow and interface—contact surfaces transmitting the load of the flow. It was necessary to simulate the blood flow for each of the vessel-valve system, and then import the designated pressure, generated by the flow, on the components of the heart valve [7]. The flow analysis of both valves required the use of the rheological model of blood which is known as non-Newtonian liquid [13]. The average density of blood was set as 1040 kg m^{-3} and the aortic blood pressure as 16 kPa.

Dynamic viscosity in the aorta was determined by scaling law, which determines the nonlinear relationship between extortion and response of the system. The character of this type occurs in non-Newtonian fluids and describes the tension- deformation speed relationship:

$$V_p = V_{max} \cdot \left| \frac{1 - r}{R_{max}} \right|^{0.143}, \tag{1}$$

$$Pl = \eta_0 \cdot (\min(UpperS, sstrnr + LowerS))^{n-1}, \tag{2}$$

with: $LowerS = 0.001 \text{ s}^{-1}$, $UpperS = 327 \text{ s}^{-1}$, $n = 0.6$, $\eta_0 = 0.035 \text{ kg m}^{-1} \text{ s}^{-1.4}$, and $V_{max} = 1.51 \text{ m s}^{-1}$, $R_{max} = 9.5 \text{ mm}$.

The analysis was conducted using FSI module in the environment of Ansys V.15, which enabled to assess the impact of blood flow to the loads occurring on the heart valve components—especially on the discs and valve housing. In the project, analysis was divided into two main stages, which focused on the blood flow through the bileaflet heart valve and with tilting disc. The analysis was carried out in the ANSYS Multiphysics 15 software using the one-way FSI simulation, which consists of CFX flow analysis and mechanical Static Structural analysis. The preparation of the models for mechanical analysis required material assignment, discretization and fixation. The same materials were assigned to both types of heart valves in order to enable comparative analysis of the loads generated by flow [26]. Pyrolytic carbon (LTI CARBON) was used as an executive material of the discs i.e. the leaflets for the bileaflet heart valve and tilting disc for the tilt-disc valve. Stellite 21 from the group of titanium alloys was assigned as an executive material of the ring for both types of heart valves: Table 1 [3, 8].

Table 1 The mechanical properties of pyrolytic carbon and titanium alloy

Parameter	Pyrolytic carbon (LTI CARBON)	Titanium alloy Stellite 21
Density (kg/m^3)	1950	8330
Young modulus E (GPa)	29	250
Poisson's ratio ν	0.33	0.3
Tensile strength Rm (MPa)	120	710
Compressive strength Rs (MPa)	550	550

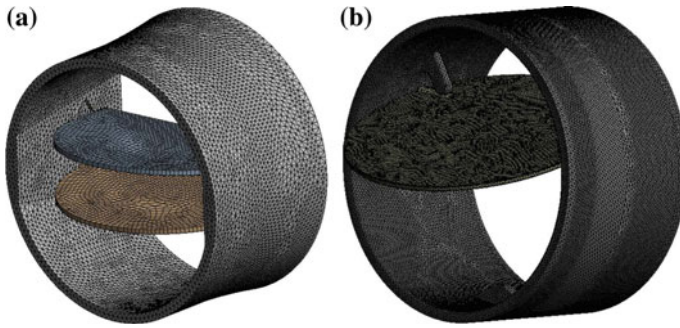


Fig. 5 Three dimensional models of the mechanical heart valves after discretization: **a** bileaflet, **b** tilting disc valve, both in open position



Fig. 6 The fixation of the models presented in the example of the bileaflet heart valve

Both models of mechanical heart valves during discretization were divided into finite elements of minimal size equal 0.3 mm (Fig. 5). Then all degrees of freedom were removed from the inlet area of the vessel (Fig. 6).

3 Results

The distribution of blood flow velocity and decrease in blood pressure is shown in Figs. 7 and 8.

In case of the bileaflet heart valve, an increase in speed was observed in relation to the opening angles. The highest velocity rate equal to 5.02 m/s was observed for closed angle in the boundary layer of the leaflets and valve housing. In these places local increase in speed was also noticeable in relation to the total flow. In the outflow part of the system turbulences have occurred probably caused by reduced flow slot. Analogously the tilting disc valve presented an increase in the flow rate depending on the opening angle. The highest rate equal to 7.44 m/s was observed for a closed angle in the boundary layer of the disc and valve housing. Local increase in speed was also observed in this area in relation to the total flow. Turbulences have occurred in the outflow part of the system probably caused by reduced flow slot. In comparison to the bileaflet heart valve, it is noticeable that flow through the valve with tilting disc is more turbulent, which may result in higher loads generation. On the other hand, the occurrence of turbulences may reduce the risk of blood clotting formed on the walls of the valve.

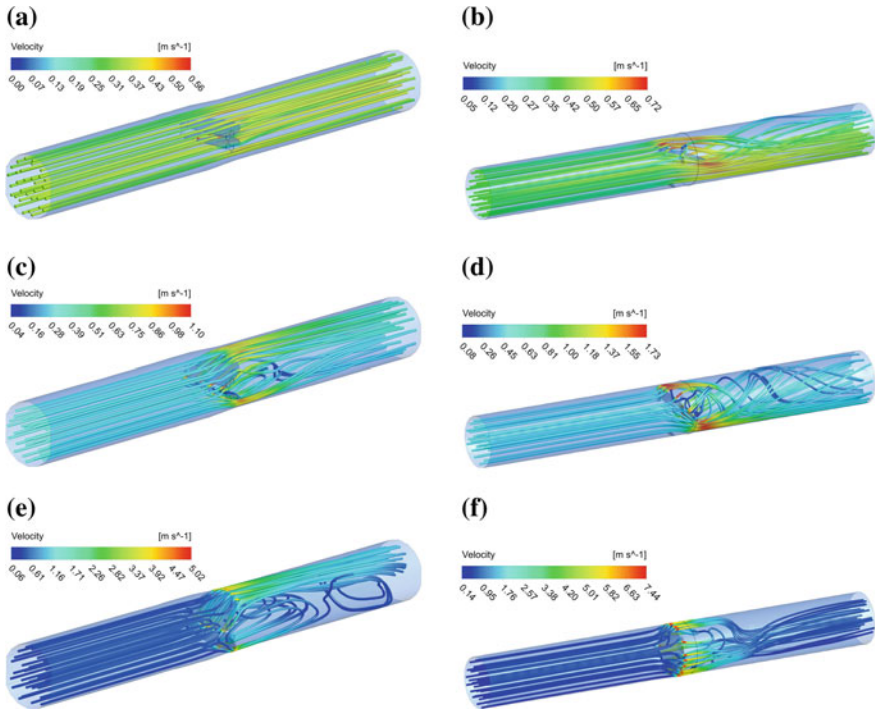


Fig. 7 The distribution of blood flow velocity through the vessel-valve system for bileaflet heart valve in open angle (a), half-open (c), closed (e) and for tilting disc valve in open angle (b), half-open (d), closed (f)

On the basis of the pressure characteristics (Fig.8), it was possible to define the maximum pressure for the bileaflet valve generated by blood, which is equal 16.17kPa for open angle. Whereas in case of the closed angle, a dramatic increase of pressure was noticed in the inflow part of the system. In the same case, the maximum pressure reaches 27.66kPa. This changes in pressure provide the force that enables propel and alternating movement of the valves.

Flow analysis for both types of valves allowed the determination of the characteristics of speed versus position of the blood in the system for all opening angles. For this purpose, authors generated the flow by point (Fig.9) and exported necessary data. The flow velocity characteristics generated by point flow for open angle (a), half-open (b) and closed angle (c) are shown in Fig. 10.

The presence of leaflets and tilt disk in the flow path has visible influence on blood flow. For both types of valves in open angle the flow is more laminar with steady flow velocity. In the case of half-open angle, a slight increase in speed was noticed at the height of the valve, which proves the significant influence of the valve on blood flow. It is also visible that a gradual increase in speed occurs in case of tilting disc valve in closed angle, which is different in comparison to bileaflet valve. A process

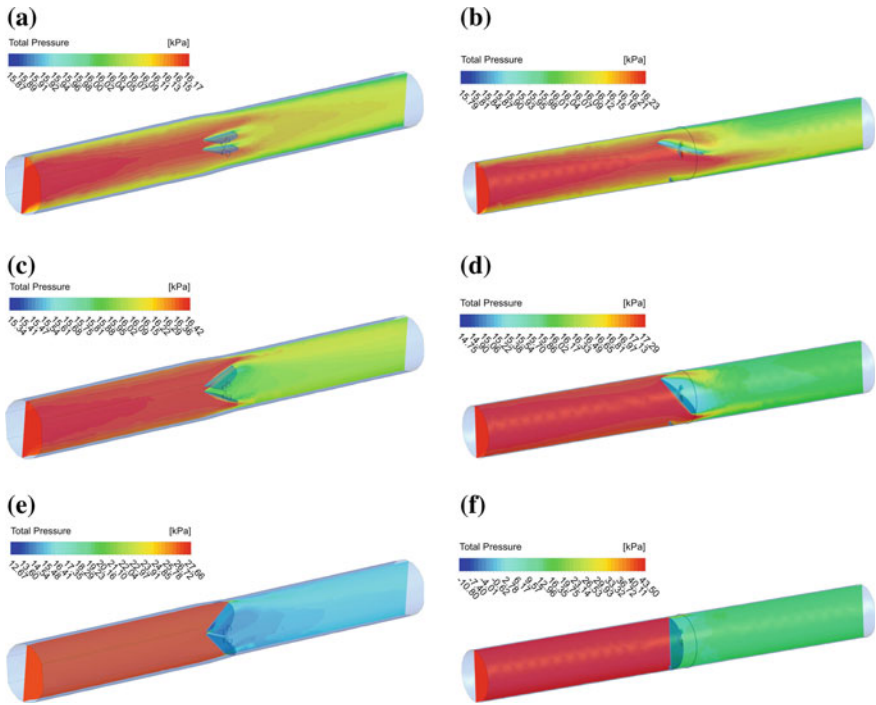


Fig. 8 The distribution of blood pressure drop during the flow through the vessel-valve system for bileaflet heart valve in open angle (a), half-open (c), closed (e) and for tilting disc valve in open angle (b), half-open (d), closed (f)



Fig. 9 The flow by point presented in the example for bileaflet heart valve

of progressive increase in flow velocity can be caused by the occurrence of local turbulences during the flow. However flow velocity gets gradually stabilized after moving barrier, which is in this case closed valve.

As a result of the CFX analysis, the pressure generated by the blood flow was obtained. On this basis it was possible to define the impact of the flow on the components of the valves. During the analysis the stress and displacement were measured for both valves in all positions (Figs. 11 and 12).

The stress was measured mainly in the outer area of the disc, which is responsible for disc positioning. This situation was observed for both heart valves in all opening angles and was a result due to the contact between the disc and the housing to enable valve movement. The maximum stress was observed for closed position and for

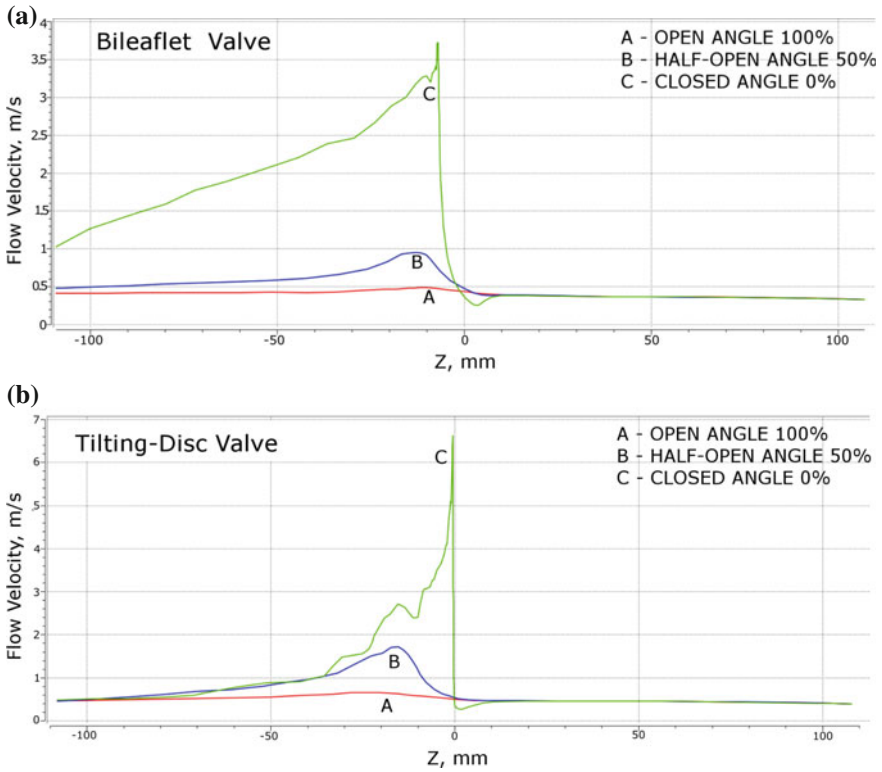


Fig. 10 The flow velocity characteristics generated by point flow for different open angles for bileaflet heart valve (a) and tilting disc valve (b)

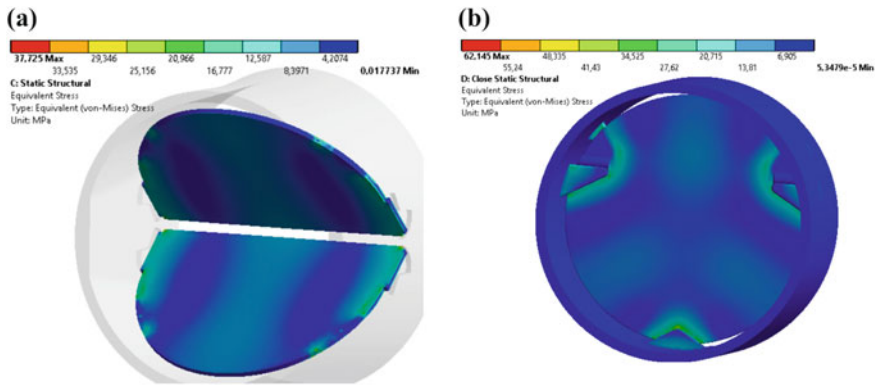


Fig. 11 The Stress distribution for heart valves exemplified with closed angle for: a bileaflet valve, b tilting disc valve

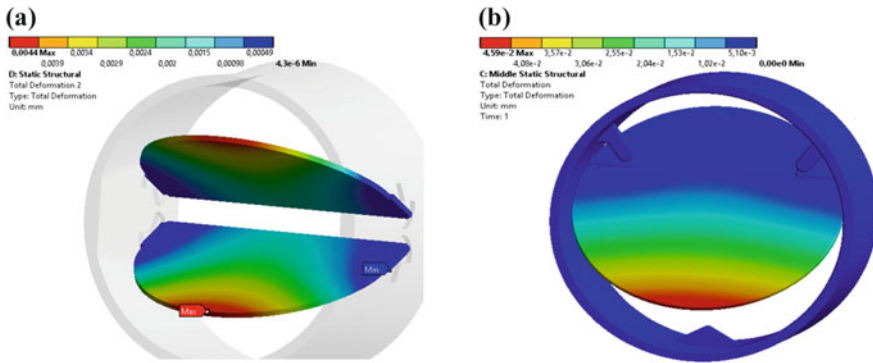


Fig. 12 The Displacement distribution for heart valves exemplified with half-open angle for: **a** bileaflet valve **b** tilting disc valve

bileaflet valve is equal 38 MPa. In comparison, on the components of the tilting disc the stress rises up to 62 MPa.

In the case of displacement, it was observed that the maximum value was located in the outer part of the disc and the leaflets. The maximum value was obtained for the bileaflet valve and is equal $43 \mu m$. However the maximum displacement of the tilting disc valve has reached $37 \mu m$.

4 Conclusion

The conducted analysis of blood flow using one-way FSI module allowed determination of the characteristics of blood pressure and flow velocity at different opening angles. The results indicate that the bileaflet valve is characterized by more laminar flow in comparison to tilting disc valve, whereas in this type of valve local turbulences were observed. However these turbulences may have a positive impact on reduction of blood clotting in such valves, which is quite advantageous. On the other hand in case of bileaflet heart valve there is a greater risk of hypertrophy and thrombogenesis. The analysis revealed that the bileaflet valve has a more steady blood flow. The maximum stress was recorded for both valves in the closed position in the center of the leaflets. The stress concentrations in this area are caused by the maximum speed that the blood stream reaches in this zone.

The maximum displacement caused by the flow was measured for a tilting disc valve in half-open position. Consequently the generated strain may cause damage to the disc during its movement. Therefore, it is recommended to use an alternative biomaterial, e.g., Polyether ether ketone (PEEK). Information obtained during such analysis may be the basis to identify alternative biomaterials for mechanical heart valves, but also allows the observation of the phenomena of blood flow.

References

1. Basiaga, M., Paszenda, Z., Szewczenko, J.: Biomechanical behaviour of surgical drills in simulated conditions of drilling in a bone. *Information Technologies in Biomedicine*. Springer, Berlin (2010)
2. Basiaga, M., Paszenda, Z., Szewczenko, J., Kaczmarek, M.: Influence of surgical drills wear on thermal process generated in the femur model. *Acta Bioeng. Biomech.* **15** (2013)
3. Belik, R.V.: *Researches in Technology of Pyrographite Items Production*, Russia (1987)
4. Bronzino, J., Park, J.: *Biomaterials Principles and Applications*. CRC Press, Boca Raton (2002)
5. Bronzino, J.: *Tissue Engineering and Artificial Organs*. CRC Press, Boca Raton (2006)
6. Butchart, E.G., Hui-Hua, L., Payne, N.: Twenty years' experience with Medtronic Hall valve. *J.Thorac. Cardiovasc. Surg.* (2001)
7. Dasi, L., Simon, H.: Fluid mechanics of artificial heart valves. *Clinical and Experimental Pharmacology and Physiology*, vol. 36 (2009)
8. Dauskardt, R.H., Ritchie, R.O.: *Pyrolytic Carbon Coatings*. World Scientific Publ.Co, Singapore (1993)
9. Enderle, J.: *Introduction to Biomedical Engineering*. Elsevier Inc. (2005)
10. Gott, V.L., Alejo, D.E., Cameron, D.E.: *Mechanical Heart Valves: 50 years of Evolution*. Division of Cardiac Surgery, Baltimore, Maryland, U.S. (2003)
11. Gudarzi, M., Zamanian, H., Oveisi, A.: A steady flow analysis of blood flow properties through some defective bileaflet mechanical heart valves. *TJEAS J.* (2013)
12. Hench, L.: *Biomaterials, Artificial Organs and Tissue Engineering*. Woodhead Publishing (2006)
13. Jozwik, K., Obidowski, D.: *Biomechanical Engineering Laboratory#3 Laboratory Instruction*, Technical University of Lodz Publishing (2010)
14. Kajzer, W., Krauze, A., Kaczmarek, M., Marciniak, J.: FEM analysis of the expandable intramedullary nail. In: *Conference on Information Technologies in Biomedicine 2008*. Springer (2008)
15. Kajzer, W., Kajzer, A., Gzik-Zroska, B., Wolanski, W., Janicka, I., Dzielicki, J.: Comparison of Numerical and Experimental Analysis of Plates Used in Treatment of Anterior Surface Deformity of Chest. *ITIB 2012*. Springer (2012)
16. Kajzer, W., Kajzer, A., Gzik-Zroska, B., Wolanski, W., Janicka, I., Dzielicki, J.: Experimental biomechanical assessment of plate stabilizers for treatment of pectus excavatum. *Acta Bioeng. Biomech.* **15** (2013)
17. Kalugin, V.I., Kozirkin, B.I.: *Electronic Industry*, Russia (1984)
18. Kiang-ia, A., Chatpun, S.: Mechanical analysis of mechanical aortic heart valve: Trileaflet versus bileaflet, *BMEiCON* (2013)
19. Kiel, M., Marciniak, J., Basiaga, M., Szewczenko, J.: Numerical Analysis of Spine Stabilizers on Lumbar Part of Spine. *Information Technologies in Biomedicine*. Springer, Berlin (2010)
20. Marciniak, J., Szewczenko, J., Walke, W., Basiaga, M., Kiel, M., Manka, I.: Biomechanical analysis of lumbar spine stabilization by means of transpedicular stabilizer. *Information Technologies in Biomedicine*. Springer (2008)
21. Pochrzast, M., Basiaga, M., Marciniak, J., Kaczmarek, M.: Biomechanical analysis of limited-contact plate used for osteosynthesis. *Acta Bioeng. Biomech.* **16** (2014)
22. Szuber, A., Kapis, A., Kustosz, R.: Polish mechanical heart valve based on MOLL's construction designed for use with pulsating heart assist pumps Religa Heart EXT. *Monograph of The Polish Artificial Heart* (2013)
23. Tullio, M.D.: Fluid dynamics and blood damage in artificial heart valves: biological vs. mechanical aortic prostheses, *WCCM XI* (2014)
24. Vahanian, A., Alfieri, O., Al-Attar, N.: Transcatheter valve implantation for patients with aortic stenosis: a position statement from EACTS and ESC, in collaboration with the EAPCI. *Eur. Heart J.* (2008)
25. Veseley, I.: Heart valve tissue engineering. *Circ. Res.* (2005)

26. Volkov, G.M., Dobrova, N.B., Zaharova, E.N.: Carbonic material for prosthetic heart valve. Col, Carbonic construction materials (1979)
27. Waite, L.: *Biofluid Mechanics in Cardiovascular System*. McGraw Education (2005)
28. Wieting, D.W.: The Bjork-Shiley Delrin tilting disc heart valve: historical perspective, design and need for scientific analyses after 25 years. *J. Heart Valve Dis.* (1996)
29. Yuan, Q., Xu, L., Ngoi, B., Yeo, T., Hwang, N.: Dynamic impact stress analysis of a bileaflet. *J. Heart Valve Dis.* (2003)
30. Ziebowicz, A., Kajzer A., Kajzer W., Marciniak J.: Metatarsal osteotomy using double-threaded screws—biomechanical analysis. In: *Conference on Information Technologies in Biomedicine 2010*. Springer, Berlin (2010)

Forming an Occlusal Splint to Support the Therapy of Bruxism

Leszek Luchowski, Agnieszka Anna Tomaka, Krzysztof Skabek,
Michał Tarnawski and Przemysław Kowalski

Abstract The present work proposes a computerised and algorithmic approach to a procedure which has traditionally been performed in hardware, by milling and polishing a physical object. While some fine physical final adjustments will still be necessary, we hope a digital approach will make the process less labor-intensive and allow for a more explicit positioning of the mandible.

Keywords Bruxism · Occlusal splint · 3D imaging · 3D prototyping

1 Principle of Operation

Bruxism is a common condition associated with nocturnal or daytime teeth clenching or grinding. Some authors classify it as a parafunction (an activity which is not related to physiology).

Although bruxism is known to be treatment-resistant, its effects (teeth wear, headaches, gingival recessions, disorders of the temporo-mandibular joint—TMJ) can be diminished. The prophylactics includes mainly behavioral (daytime bruxism) and splint therapy (both nocturnal and daytime bruxism) [4, 6]. The concept of treating bruxism with an occlusal splint relies on a reflexive path: proprioceptive receptors detecting pressure on the front teeth (canines and incisors) cause mastication muscles to relax, while a similar stimulus on the back teeth (premolars and molars) stimulates them to contract. The splint should be formed in such a way that grinding or clenching of teeth will cause equilibrated contact of the anterior teeth, triggering a relaxation [6].

L. Luchowski (✉) · A.A. Tomaka · K. Skabek · M. Tarnawski · P. Kowalski
Polish Academy of Sciences, Institute of Theoretical
and Applied Informatics, Gliwice, Poland
e-mail: leszek.luchowski@iitis.pl
URL: <http://www.iitis.pl>

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_24

Graber et al. [5] point to more reasons why occlusal devices in general reduce various symptoms related to temporo-mandibular joint disorders (TMD). Those of them that are relevant in the context of bruxism can be summarized as:

1. adjusting the position of the mandible to one that:
 - increases the vertical direction.
 - is more comfortable (optimal, stable) both from an occlusal and neuromuscular point of view.all of which factors help reduce muscular activity.
2. influencing the central nervous system (CNS):
 - the device acting as a reminder for the patient to consciously avoid undesirable behavior.
 - inhibiting the CNS during sleep, reducing muscle hyperactivity.
 - creating a placebo effect.

2 Existing Procedure

The splint is initially formed with an impression of the upper teeth on one side and only of the cusps of the lower teeth on the other. The lower surface is then cut, by a technician, in the areas surrounding the cusp impressions so as to reduce it to nearly a plane, where the points corresponding to cusp tips (which remain intact) are no longer depressed. This near-plane is finished by the doctor, in the patient's mouth, with the use of marking paper, to ensure that only the cusps and edges of lower teeth create contact points with the splint. Then, again with the use of marking paper, the patient is asked to move the mandible somewhat sideways and forward. Then, with a marking paper of another (second) color, the patient taps the teeth against the splint (in the neutral position of initial contacts). The splint is then reduced along the lines that were drawn on the lateral teeth (premolars and molars), while avoiding the points marked in the second color. Certain lines (drawn by canines and incisors) are left intact on the front part of the splint, namely:

- lines representing lateral movements of the canines are left on the side of the movement direction.
- lines representing movement forward [protrusion] are left in the area corresponding to incisors and/or canines [6].

There is some disagreement whether the incisors should be in contact during protrusion [1]. The effect is that lateral and frontal movements of the mandible, typical of grinding, are countered by the splint pressing against the front teeth (canines and incisors), causing the masticatory muscles to relax. The back teeth (premolars and molars) do not come under pressure, so that their proprioceptors—which would stimulate the masticatory muscles—are not activated.

We propose a partly computerized method which seems likely to achieve the same physiological effect while using less manual cutting and operating in a coordinate system related to CBCT (cone beam computed tomography) data.

3 Outline of the Method

The principal idea of the method is to duplicate the occlusal surface of the maxillary teeth and use its shape in two ways: in the exact negative, to form an upper surface of the splint to tightly match the maxillary teeth, and in a modified positive, to form the lower surface to accommodate the mandibular teeth in a natural way while allowing some motion and selectively providing support in the areas where it is desirable from the therapeutic point of view.

The solid of the splint is completed by adding an arbitrary smooth surface on the lingual and labial/buccal side.

A more detailed description of the proposed method is given in the following pseudocode:

1. Forming the upper surface:
 - (a) taking an impression of the upper dental arch, casting a positive model and scanning it to obtain a 3D digital representation.
 - (b) subtracting (in set-theoretic sense) this volume from a predefined, horse-shoe-shaped solid slightly wider than the dental arch.
 - (c) increasing the contact area by morphological post-processing operations such as opening and closing.
2. Forming an initial position of the lower surface of the splint by moving (virtually) a copy of the upper surface downwards by a predefined distance d (Fig. 1). This distance becomes the initial thickness of the splint.
3. Determining the contact points between this surface and the lower dental arch during jaw closure (algorithm described by Skabek and Tomaka [7]) and the corresponding position of the TMJ resulting from the initial thickness.
4. Determining the desirable position of the mandible by manipulating its 3D CBCT model, in a virtual setting, placing the condyles against the fossa and disc of the temporo-mandibular joint. The position, optimized for rotation-point centering and right-left symmetry identical to Centric Relation [4], is expressed as a rigid-body transformation T (rotation and translation) from the position obtained in step 2.
5. Forming the lower surface of the splint by:
 - (a) taking the copy of the upper dental arch shifted in step 2.
 - (b) moving it further by applying T —Fig. 1.
 - (c) adding (in a set-theoretic sense) the resulting volume to the solid created in step 1c. whose thickness may have to be increased to avoid gaps.
 - (d) filling and smoothing the edges created at the joining.

6. Prototyping the splint in a 3D printer. If a printer capable of working with biocompatible materials is not available, prototyping a mold in which a medical-grade splint will be cast.
7. Finishing the splint by using marking paper (in two different colors) and a dental bur as in the existing procedure described at the beginning of this paper [3, 7] (Fig. 2).

Fig. 1 Moving a copy of the lower surface of the maxillary arch. **a** Natural occlusion, **b** vertical shift, **c** general rigid-body movement

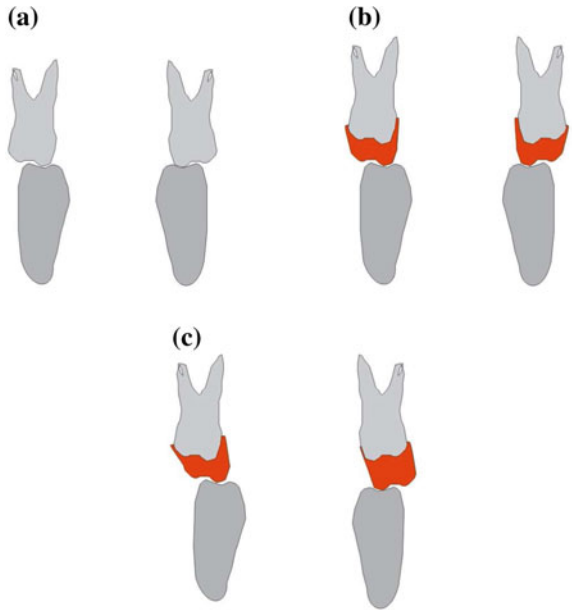
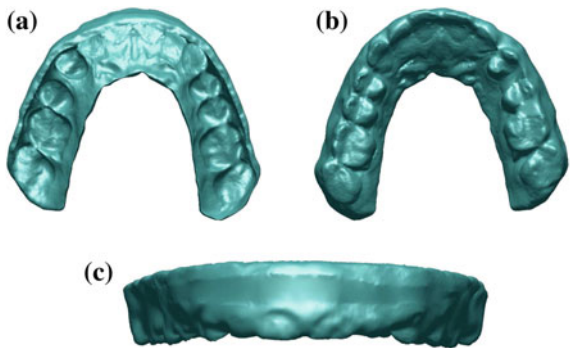


Fig. 2 Draft of occlusal splint. **a** Top side shape of maxillary teeth in negative, **b** bottom side shape of maxillary teeth in positive, **c** frontal view showing labial and buccal surfaces



3.1 User Interface Issues

In point 4 above, we mentioned a virtual setting in which the doctor should manipulate a CBCT model of the mandible, seeking its optimal position from a therapeutic point of view. Although various application programs are available on the market that allow the user to position two solids relative to each other, doing so with the mandible and maxilla requires some dedicated verifying functionalities. The interface should allow the doctor to manually modify the position of the mandible relative to the maxilla by small increments, while verifying, at each increment, both the occlusal and condylar conditions. Small increments here mean rotation by small angles and translation by small vectors. Occlusal conditions—the contact between the dental arches—can be verified by analyzing the cross-sections of the interacting solids, while condylar relations (position of the TMJ) can be visualized by a virtual X-ray projection [8]. The use of virtual X-ray mirrors the long practice of evaluating the position of the TMJ by means of 2D X-ray images. 2D has long been the norm in medical imaging in general, and many experienced practitioners can interpret this type of images better than a perspective or axonometric projection. A number of criteria which the practitioner will use when evaluating the desirable position of the mandible are also defined in 2D terms. While future implementations can obviously use 3D screens or goggles as well as 3D pointing devices, X-ray mode visualization is likely to remain a desirable option.

After each adjustment of the position, the system should check for any collisions, to rule out any move that would make the two solids (the maxilla and the mandible) intersect.

As an illustration of concept only, Fig. 3 is a screenshot of a model of the mandible being moved by means of a virtual trackball. It was taken under RapidForm (by InUs Technologies), a general-purpose 3D editor which does not automatically prevent

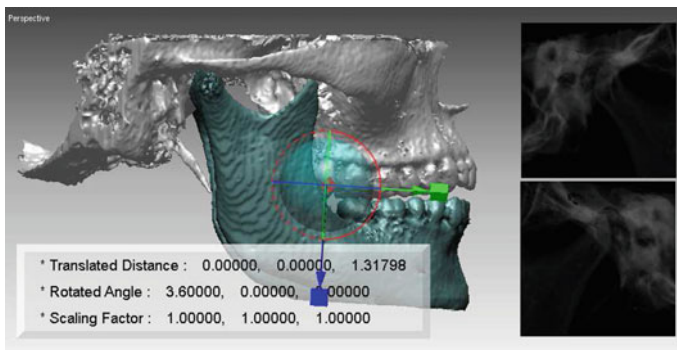
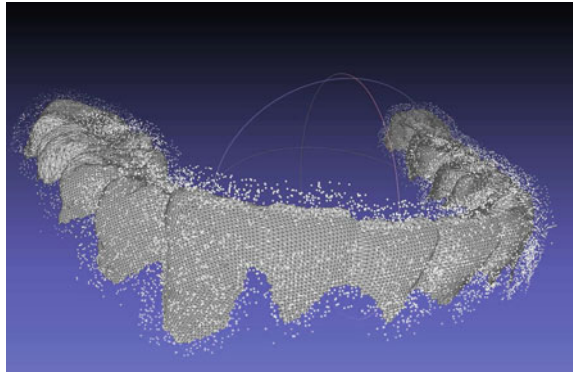


Fig. 3 Moving the mandible in a virtual environment

Fig. 4 Thickened model of the mandibular dental arch



collisions. The procedure uses a standard 2D computer screen and mouse, and the user can switch between several modes of using the mouse:

- translational movement in either the XY, YZ, or XZ plane;
- rotation around the X, Y or Z axis.

4 Virtual Finishing

A part of the physical finishing process might be replaced by mathematical morphology operations such as thickening and thinning, performed in the virtual domain. These operations are similar to erosion and dilation (from mathematical morphology) but are easier to implement for the 3D meshes we are working with, whereas erosion and dilation are better suited for volumetric data.

In a first version, our algorithm only moves the vertices—the topology of the mesh remains unchanged. As the direction of the movement, we use a pseudo-normal vector calculated using adjacent facets. The normal vector is calculated as a weighted average, where we use the adjacent angle as a weight for face normal (the method for calculating pseudo-normal is described in [2]). Results are shown in Fig. 4.

Acknowledgments This work was supported in part by the Project 2012/07 /B/ST6/01238, sponsored by the NCN (the Polish National Center of Science).

References

1. Ash, M.M. Jr., Ramfjord, S.P., Schmidseder, J.: *Terapia przy użyciu szyn okluzyjnych (Occlusal splints in therapy)*. Wyd. polskie red. Teresa Maślanka, Urban & partner, Wrocław (1999) (in Polish)
2. Baerentzen, J.A., Aanaes, H.: Signed distance computation using the angle weighted pseudo-normal. *IEEE Trans. Vis. Comput. Graph.* **11**(3), 243–253 (2005)

3. Baron, S.: Bruxism and Occlusion. In: Palla, S., Kato, T., Provaznik, I., Tkacz, E., Baron, S. (eds) *Novel Methodology of Both Diagnosis and Therapy of Bruxism, Lecture Notes of the ICB Seminar, 137th Seminar, Warsaw (2014)*
4. Dawson, P.E.: *Functional Occlusion. From TMJ to Smile Design*. Mosby-Elsevier, St. Louis (2007)
5. Graber, L.W., Vanarsdall, R.L., Vig, K.W.L.: *Orthodontics: Current Principles and Techniques*. Elsevier Health Sciences (2011)
6. Okeson, J.P.: Leczenie dysfunkcji narządu żucia i zaburzeń zwarcia (Treatment of masticatory disfunction and malocclusion). In: Drobek, W. (ed) *Wydawnictwo Czelej, Lublin (2005)*
7. Skabek, K., Tomaka, A.A.: Techniques of processing and segmentation of 3D geometric models for detection of the occlusal area of dental arches. In: Palla, S., Kato, T., Provaznik, I., Tkacz, E., Baron, S. (eds) *Novel Methodology of Both Diagnosis and Therapy of Bruxism, Lecture Notes of the ICB Seminar, 137th Seminar, Warsaw (2014)*
8. Tomaka, A.A.: Analiza obrazów wielomodalnych dla potrzeb nieinwazyjnej diagnostyki ortodontycznej (Multimodal image analysis for noninvasive orthodontic diagnosis). Instytut Informatyki Teoretycznej i Stosowanej PAN, D.Sc. thesis (2013) (in Polish)

Sensitivity Analysis of Signaling Pathways in the Frequency Domain

Malgorzata Kardynska and Jaroslaw Smieja

Abstract Sensitivity analysis methods have been in use for over half a century. In recent years they have been applied also to perform analysis of biological systems models. While sensitivity methods proved to be helpful in analysis of various pathway models whose simulation results units were fixed, they may lead to false conclusions when the model is based on relative data. For this reason it is necessary to develop methods which take into account specificity of such experimental data. In this paper we propose a method based on frequency distribution of a system time response to create parameter rankings, that can be used to find the most important parameters and processes and subsequently can be used e.g. to indicate molecular targets for new drugs. In order to test the proposed procedure we performed sensitivity analysis of several signaling pathways and we compared the results with the sensitivity functions method.

Keywords Sensitivity analysis · Signaling pathways · Fourier frequency analysis · p53 · Interferon- β

1 Introduction

Dynamic computational models play an increasingly important role in systems biology. Such models are powerful tools that allow us to develop and test several hypotheses about complex biological systems [6, 13]. An example of such a system can be signaling pathways—cascades of processes, which may be initiated by an external event (e.g., ligand binding to its specific receptor on a cell surface), or by an internal

M. Kardynska (✉) · J. Smieja
Faculty of Automatic Control, Electronics and Computer Science,
Institute of Automatic Control, Silesian University of Technology,
Akademicka 16, Gliwice, Poland
e-mail: malgorzata.kardynska@polsl.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_25

275

event (e.g., DNA damage). In the literature there is a growing number of high dimensional models with a large number of parameters. However, methods for measuring biochemical parameters are limited and may introduce substantial inaccuracies [7]. Therefore, each model should be checked with respect to its sensitivity to parameter changes.

In this paper we focused on models given by ordinary differential equations that describe change in concentration of molecules involved in a signaling pathway. Such models should be robust with respect to changes in parameters in a relatively wide range, which may represent the differences between individual cells. Sensitivity analysis of those models will not only validate the model, but also provide information about the most important processes and parameters of the model, which should be determined with the highest accuracy.

Sensitivity analysis methods have been used to test mathematical models for over half a century. However, the methods used e.g. in automatic control cannot always be directly used in systems biology. It is necessary to develop methods which take into account the specificity of biological systems and experimental data. In many cases biological experiments provide data about the fold increase of the number of molecules or of the concentration, while their absolute values are not known. It has been shown that the methods of sensitivity analysis, which have been proven to be a useful tool to analyze the quantitative models [1, 2, 9, 10, 14], may lead to false conclusions when the model is based on such relative data [12].

In this paper we compare two methods of sensitivity analysis: (i) sensitivity functions [5], which are one of the simplest, but still widely used method, and (ii) our original method of sensitivity analysis in the frequency domain. In order to compare these two methods we analyzed two signaling pathways with varying complexity: a toy model of p53 signaling pathway [3] and a complex model of interferon- β induced signaling pathway [11].

2 Methods

Sensitivity analysis is usually divided into two major categories: local and global. Local sensitivity analysis describes how the system output changes when parameter values deviate in a small range. Global sensitivities, in turn, describe how the system output changes when multiple parameters change in a relatively wide range [4, 5] and allow to capture general sensitivity of the system, however usually they require much more computational power. In this paper we focused only on local methods of sensitivity analysis. We present our method of sensitivity analysis based on frequency distribution obtained by its Fourier transform and compare it with the standard method of sensitivity functions.

2.1 Sensitivity Functions

The model represents a reaction network, described by a nonlinear state equation:

$$\frac{dX}{dt} = f(X, u, \theta), \quad (1)$$

where $X = [x_1, x_2, \dots, x_n]^T$ is a state vector with x_i denoting concentration of molecules of type i , u is an input variable, which in the context of signaling pathways may be either step or impulse excitations, and θ are model parameters. The sensitivity functions s_{ij} , describing the influence of the j -th parameter on the i -th state variable are defined as:

$$s_{ij} = \frac{\partial x_i}{\partial \theta_j}, \quad (2)$$

where x_i denotes the i -th state variable and θ_j denotes the j -th parameter, and the absolute sensitivity matrix is defined as:

$$S = \frac{\partial X}{\partial \theta}. \quad (3)$$

Since an analytical solution of (1) is usually not available, sensitivity coefficients must be calculated in some other way. One of the approaches that can be applied is the direct differential method [5]. Calculating partial derivative of (1) with respect to θ_j yields:

$$\frac{d}{dt} \frac{\partial X}{\partial \theta_j} = \frac{\partial f}{\partial X} \frac{\partial X}{\partial \theta_j} + \frac{\partial f}{\partial \theta_j} = J \cdot S_j + F_j, \quad (4)$$

where J is the Jacobian matrix, F_j is the parametric Jacobian matrix, and S_j is the column sensitivity vector with respect to the j -th parameter.

In biological systems different parameters may take values that are distributed over several orders of magnitude. The same holds true also for model state variables. Therefore, instead of the absolute sensitivities, the relative sensitivities are:

$$s_{ij} = \frac{\partial x_i}{\partial \theta_j} \frac{\theta_j}{x_i}. \quad (5)$$

Such normalization makes it possible to compare relative influence of any parameter change on system behavior, regardless of the scale of either the parameter or the state variable.

Parameter rankings that can be based, among other things, on sensitivity functions, provide a useful way to compare importance of parameters. They can be calculated either for each state variable separately, or for the whole system. In this paper we

focused on parameters rankings for individual variables. The importance of the j -th parameter for the i -th state variable can be measured as:

$$RS_{i,j}^* = \frac{1}{T} \int_0^T |s_{i,j}(\tau)| d\tau, \quad (6)$$

where T denotes the time horizon of the simulation. Parameter rankings can be used to find the most important parameters and processes (that have the greatest impact on the system output) and subsequently can be used to reduce the number of experiments needed to precisely estimate parameters values or to indicate molecular targets for new drugs.

2.2 Fourier Sensitivity Analysis

Here we present a new method for creating parameters rankings based on frequency distribution of a system time response, which can be obtained using Discrete Fourier Transform (DFT). The general algorithm for the model described by (1) is defined by the following steps:

1. Run the simulation for nominal parameter values θ_n , obtaining $X_{ref}(t, \theta_n)$.
2. Calculate Discrete Fourier Transform (DFT) of $x_{i,ref}(t, \theta_n)$, denoted by $F_{i,ref}(\omega, \theta_n)$ (here, we concentrate on a single state variable x_i).
3. For each parameter, generate a large set of its random values from a distribution determined by what is known about biochemically feasible parameter values.
4. For each generated parameter value:
 - run a simulation with remaining parameters at their nominal values $x_i(t, \theta_j)$,
 - calculate the difference between the nominal response and the new response, defined as:

$$\Delta x_i(t, \theta_j) = x_{i,ref}(t, \theta_n) - x_i(t, \theta_j), \quad (7)$$

where x_i denotes concentration of variable i , θ_n is a nominal parameters vector, and θ_j is new parameters vector with one parameter replaced by randomly generated value,

- calculate Discrete Fourier Transform of $\Delta x_i(t, \theta_j)$ denoted by $\Delta F_i(\omega, \theta_j)$,
- calculate the index of deviation of the frequency response given by [8]:

$$AA_j = \frac{\sum_{n=0}^{2^m} |\Delta F_i(\omega, \theta_j)|}{\sum_{n=0}^{2^m} |F_{i,ref}(\omega, \theta_n)|}, \quad (8)$$

where $\Delta F_i(\omega, \theta_j)$ is the difference signal amplitude at frequency ω and $F_{i,ref}(\omega, \theta_n)$ is the reference signal amplitude at frequency ω .

5. Calculate the mean and variance of the AA index for each parameter and use it to determine its Fano factor, subsequently used in parameter ranking:

$$\hat{s}_{ij} = \frac{\sigma^2(AA_j)}{\mu(AA_j)}, \quad (9)$$

where $\sigma^2(AA_j)$ and $\mu(AA_j)$ are variance and mean of the mean magnitude of $F_i(\omega, \theta_j)$ calculated for simulations in which θ_j was varied.

Usually, for signaling pathways models parameter values are randomly generated from one of two possible distributions:

- normal or Gamma distribution, if the nominal value of a parameter has already been determined or is known from literature,
- uniform distribution, defined on a wide range of biologically acceptable values, if the parameters are not known.

Since frequency distribution provides valuable information about qualitative and quantitative system behavior it can be used to analyze both: qualitative and quantitative models. This is very useful for biological models, which are dependent on the data used to estimate the parameters. When the model parameters are estimated with use of absolute measurements, the standard version of the algorithm can be used. However, many experimental methods (e.g. microarrays, Real-Time PCR, electrophoretic gel assays, etc.) do not provide measurements in absolute concentrations. Instead the level of a given protein, protein complex or mRNA is given as relative to some control (usually, the level of the same molecule type at time 0). In this case some kind of normalization must be done in order to reduce the influence of the steady state, whose value cannot be measured. We propose the normalization with respect to the area under curve of system time responses calculated for each variable separately. The normalization should be implemented after run of every simulation (before computing the Fourier transform).

3 Results and Discussion

In this section results of sensitivity analysis of two signaling pathways models: toy model of p53 signaling pathway and a complex model of interferon- β induced signaling pathway, are presented. The first one is much simpler, making it easier to understand the model and validate the results of sensitivity analysis. However, such simple models are rare in the case of signaling pathways. More frequently we see very complex models containing dozens of state variables and multiple positive and negative feedback loops. An example of such a complex model is the interferon- β induced signaling pathway.

3.1 P53 Pathway Toy Model

The dynamics of p53/Mdm2 regulatory module under consideration is described fully in [3]. It is a relatively simple model including three state variables: total p53 protein, cytoplasmic Mdm2 and nuclear Mdm2, and 10 parameters listed in Table 1.

The parameters m and n are the numbers of p53 and Mdm2 gene copies, respectively, s_1, s_2, s_3 and s_4 are the production rates per gene copy, k_{d1} and k_{d2} are p53 and Mdm2 degradation rates and k_1, k_2 are Mdm2-mediated nuclear import rates. Since a full sensitivity analysis of the model is beyond the scope of this paper, further discussion will take into account only the sensitivity analysis of the concentration of total p53 protein.

Following the procedure described in the previous section we investigated the p53 variable sensitivity to parameter changes. The parameters m and n were omitted in the analysis, because in normal cells their values are known and unchanging. The parameters ranking based on sensitivity functions method is shown in Fig. 1a, while parameters rankings based on frequency distribution are shown in Fig. 1b, c for normal and normalized system time responses, respectively.

Table 1 Simple p53/Mdm2 model parameters

No.	Symbol	Value	No.	Symbol	Value
1	s_1	16	6	k_{d2}	2.2×10^{-4}
2	s_2	8	7	k_1	3.5×10^{-3}
3	s_3	80	8	k_2	2300
4	s_4	105	9	m	2
5	k_{d1}	10213	10	n	2

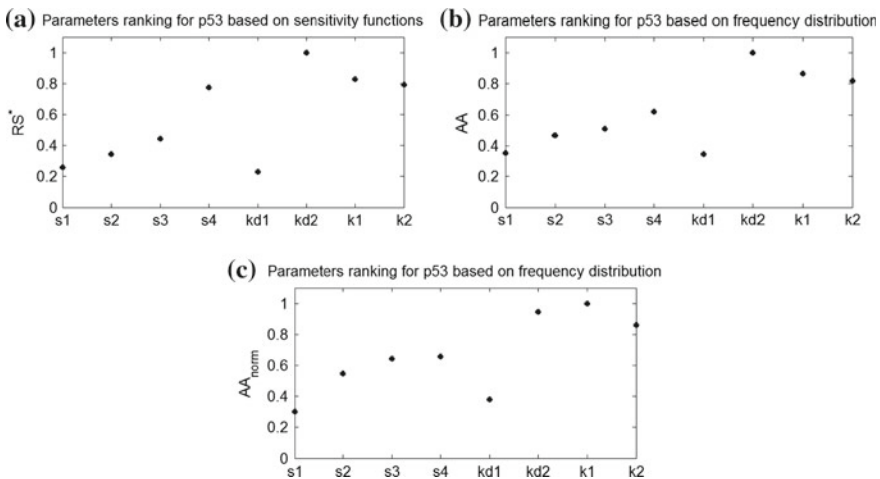


Fig. 1 Parameter rankings for p53 based on: sensitivity functions (a), frequency distribution (b), frequency distribution with normalization (c)

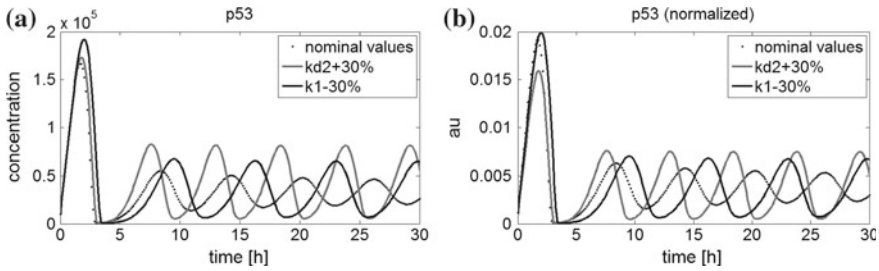


Fig. 2 The comparison of p53 time responses for three simulation runs: for nominal parameter values, for parameter k_{d2} increased by 30% and for parameter k_1 decreased by 30%, without normalization (a) and with normalization (b)

In first two rankings the position of most parameters is comparable and is consistent with the results of the simulation analysis. Differences appear only after adding the normalization to the proposed algorithm of sensitivity analysis in the frequency domain. In this case, the influence of the parameter k_1 slightly increases. To illustrate it we performed two set of simulations: with and without normalization, for nominal parameter values, parameter k_{d1} increased by 30% and parameter k_1 decreased by 30%. The change direction of parameters was selected on the basis of the sensitivity functions. The results of these simulations is shown in Fig. 2.

As we can see changing the parameter k_{d2} increases the maximum concentration of p53 and results in phase shift in system response, while changing the parameter k_1 in addition changes also the frequency of oscillation. After normalization and removing the effect of amplification the impact of parameter k_1 is more visible.

3.2 Interferon- β Induced Pathway Model

The second model under consideration is a model of interferon- β induced JAK-STAT signaling pathways, described in [11]. The model is far more complex than the previous example and includes 31 state variables and 58 parameters (listed in Table 2).

We have chosen IRF1 mRNA as one of the state variables to illustrate and compare proposed sensitivity analysis methods. In Figs. 3, 4 and 5 are shown parameters rankings for chosen state variable. Due to high number of parameters the horizontal axes contains only numbers corresponding to the parameters listed in Table 2.

We can notice a significant differences in obtained parameters rankings, e.g. in parameters 27 (*kinvs1s2n*) and 53 (*ks1phos*). In both rankings based on frequency distribution parameter *kinvs1s2n* seems to be the most important, while in ranking based on sensitivity function the importance of this parameter is negligible and the parameter *ks1phos* is indicated as one of the most important. To check which of these two parameters has greater influence on the system response, we performed

Table 2 The list of parameters of interferon- β induced pathway model

No.	Symbol	Value	No.	Symbol	Value
1	kv	5×10^{-1}	30	$ks1s1pdeg$	0
2	$vtranscription$	0.5×10^{-5}	31	$ks1s2pdeg$	0
3	$vstat1transcription$	0.6×10^{-5}	32	$ks1i1deg$	2.88×10^{-5}
4	$vstat2transcription$	0	33	$kdegpiass1s1$	0
5	$vtmp2transcription$	0.2×10^{-5}	34	$ks1s1$	0.4×10^1
6	$vtap1transcription$	0.2×10^{-5}	35	$ks1s2$	1×10^1
7	$ktransl$	5×10^{-1}	36	$kpiass1s1$	7×10^{-1}
8	$constant1$	3.27×10^{-4}	37	$ks1i1$	0.1×10^{-1}
9	$constant2$	3.27×10^{-4}	38	$kactivation$	0.7×10^{-3}
10	$constant3$	3.27×10^{-4}	39	$kinacti1$	7.7×10^{-5}
11	$constant4$	3.27×10^{-4}	40	$ks1tprod$	2.73×10^{-9}
12	$constanttmp$	3.33×10^{-3}	41	$ks2tprod$	1.82×10^{-9}
13	$ks1deg$	1.28×10^{-5}	42	$tmp2tprod$	1.42×10^{-9}
14	$ks1pdeg$	1.28×10^{-5}	43	$tap1tprod$	2.13×10^{-9}
15	$ks2deg$	1.28×10^{-5}	44	$es1$	1.44×10^{-3}
16	$ks2pdeg$	1.28×10^{-5}	45	$is1$	8.16×10^{-5}
17	$ki1deg$	5.77×10^{-5}	46	$es2$	1.44×10^{-3}
18	$ki1indeg$	5.77×10^{-5}	47	$is2$	4.89×10^{-4}
19	$kdegs1t$	3.85×10^{-4}	48	$is1s1$	2.31×10^{-2}
20	$kdegs2t$	3.85×10^{-4}	49	$is1s2$	2.31×10^{-2}
21	$kdegi1t$	6.41×10^{-5}	50	$ii1$	2.31×10^{-2}
22	$kdeglmp2t$	6.41×10^{-5}	51	$ei1$	2.31×10^{-6}
23	$kdegtap1t$	9.62×10^{-5}	52	$ei1in$	5.77×10^{-3}
24	$kinvs1s1$	4.62×10^{-4}	53	$ks1phos$	8.9×10^{-1}
25	$kinvs1s1n$	4.62×10^{-4}	54	$ks1dephc$	5.77×10^{-4}
26	$kinvs1s2$	4.62×10^{-4}	55	$ks2phos$	4.2×10^{-1}
27	$kinvs1s2n$	4.62×10^{-4}	56	$ks2dephc$	5×10^{-4}
28	$kinvpiass1s1$	3.85×10^{-3}	57	$ks1phossat$	1×10^4
29	$kinvs1i1$	2.88×10^{-3}	58	$ks2phossat$	1×10^4

three simulations: for nominal parameter values, for parameter $kinvs1s2n$ decreased by 30% and for parameter $ks1phos$ decreased by 30%. The time responses of IRF1 mRNA for these three simulations are shown in Fig. 6a. For comparison, the right panel (b) presents the results of the three simulations after normalization with respect to the area under the curve of system time responses. By comparing these three time courses, we can see that both parameters significantly change the IRF1 mRNA time response and should be placed high in the parameters ranking. However, after normalization the impact of parameter $ks1phos$ significantly decreases, which corresponds to the parameters ranking in Fig. 5.

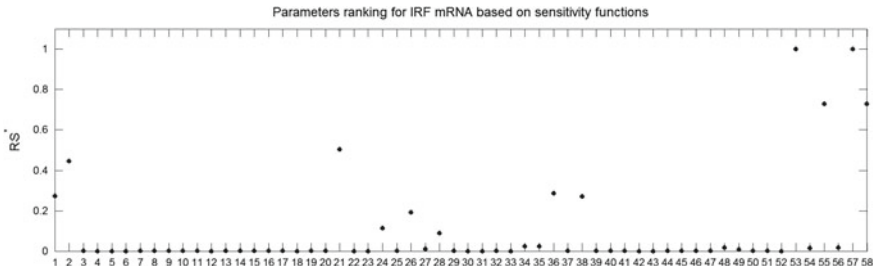


Fig. 3 Parameters ranking for IRF1 mRNA based on sensitivity functions method

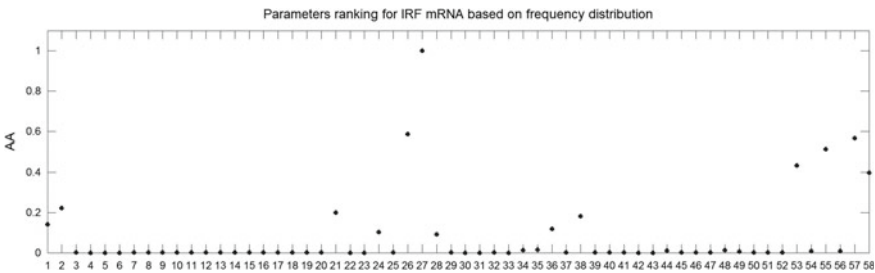


Fig. 4 Parameters ranking for IRF1 mRNA based on frequency distribution

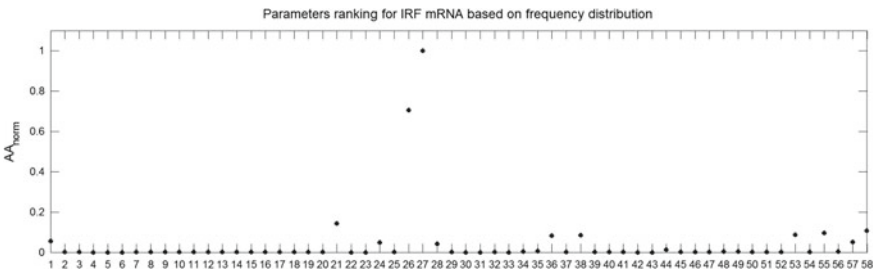


Fig. 5 Parameters ranking for normalized IRF1 mRNA based on frequency distribution

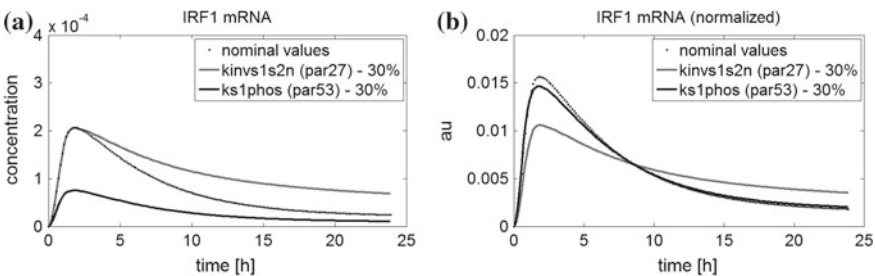


Fig. 6 The comparison of IRF1 mRNA time responses for: nominal parameter values, parameter *kinvs1s2n* decreased by 30% and parameter *ks1phos* decreased by 30%, without normalization (a) and with normalization (b)

Table 3 The results comparison of three methods: (A) Sensitivity functions, (B) Fourier sensitivity analysis, (C) Fourier sensitivity analysis with normalization

	A	B	C
3 Most important parameters	53; 57; 58	27; 26; 57	27; 26; 21
Cumulative change in the model response	4.66	5.16	–
Cumulative change in the normalized model response	0.07	–	0.37

To make it easier to compare both methods of sensitivity analysis we summarized the results in Table 3. We quantified the cumulative change in the model response as the surface area between the system response for nominal parameters and the system response when the most important parameter (according to a particular method) is changed by 30 %.

4 Conclusions

In this paper we presented a new method for creating the parameters ranking based on frequency distribution of a system time response. Taking into account the examples presented in this work we showed that the method can provide valuable information about the most important parameters that have the greatest impact on the system output. Moreover, the proposed method allows us to perform sensitivity analysis of both quantitative (standard algorithm) and qualitative (algorithm with normalization) biological models.

Acknowledgments The work has been supported by the NCN grants DEC-2013/11/B/ST7/01713 (JS) and BKM-514/RAU1/2015/t.14 (MK).

References

1. Balsa-Canto, E., Alonso, A.A., Banga, J.R.: An iterative identification procedure for dynamic modeling of biochemical networks. *BMC Sys Biol.* **4**, 11 (2010)
2. Bentele, M., Lavrik, I., Ulrich, M., Stosser, S., Heermann, D.W., Kaltho, H., Krammer, P.H., Eils, R.: Mathematical modeling reveals threshold mechanism in cd95-induced apoptosis. *The Journal of Cell Biology.* **166**(6), 839–851 (2004)
3. Hat, B., Puszynski, K., Lipniacki, T.: Exploring mechanisms of oscillations in p53 and nuclear factor- κ B systems. *IET Systems Biology.* **3**, 342–355 (2009)
4. Iglesias P., Ingalls B. (eds): *Control theory and systems biology.* MIT Press (2010)
5. Leis, J., Kramer, M.: Sensitivity analysis of systems of differential and algebraic equations. *Computers and Chemical Engineering.* **9**, 93–96 (1985)
6. Locke, J.C.W., Southern, M.M., Kozma-Bognar, L., Hibberd, V., Brown, P.E., Turner, M.S., Millar, A.J.: Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol Syst Biol.* **1**, 0013 (2005)

7. Maerkl, S.J., Quake, S.R.: A systems approach to measuring the binding energy landscapes of transcription factors. *Science*. **315**, 233–237 (2007)
8. Prosek, A., Leskovar, M., Mavko, B.: Quantitative assessment with improved fast Fourier transform based method by signal mirroring. *Nucl Eng Des*. **238**(10), 2668–2677 (2008)
9. Raia, V., Schilling, M., Boehm, M.: Dynamic mathematical modeling of IL13-induced signaling in Hodgkin and primary mediastinal B-cell lymphoma allows prediction of therapeutic targets. *Cancer Res*. **71**, 693–704 (2011)
10. Rateitschak, K., et al.: Parameter Identifiability and Sensitivity Analysis Predict Targets for Enhancement of STAT1 Activity in Pancreatic Cancer and Stellate Cells. *PLoS Comput Biol*. **8**(12), e1002815 (2012)
11. Smieja, J., Jamaluddin, M., Brasier, A.R., Kimmel, M.: Model-based analysis of interferon- β induced signaling pathway. *Bioinformatics*. **24**, 2363–2369 (2008)
12. Smieja, J., Kardynska, M., Jamroz, A.: The meaning of sensitivity functions in signaling pathways analysis. *DCDS-series B*. **10**, 2697–2707 (2014)
13. Voit, E., Neves, A.R., Santos, H.: The intricate side of systems biology. *PNAS*. **103**, 9452–9457 (2006)
14. Yue, H., Brown, M., Knowles, J., Wang, H., Broomhead, D.S., Kell, D.B.: Insights into the behaviour of systems biology models from dynamic sensitivity and identifiability analysis: a case study of an nf- κ b signalling pathway. *Molecular BioSystems*. **2**(12), 640–649 (2006)

Part IV
Biomechatronics

Development of New Testing Method of Mechanical Properties of Porcine Coronary Arteries

Bożena Gzik-Zroska, Kamil Jozzko, Wojciech Wolański
and Marek Gzik

Abstract This study aims to develop a testing methodology enabling the determination of mechanical properties of coronary arteries based on the example of porcine arteries using the MTS Insight 2 strength testing machine and digital image correlation (DIC). The conducted ballooning strength tests and static tensile tests revealed linear dependence between longitudinal and transverse strain of the arteries. The test results presented enable the evaluation and identification of the mechanical properties of porcine coronary arteries as well as prediction of the risk of the coronary artery rupture in the planned angioplasty and stent implantation procedures.

Keywords Mechanical properties · Coronary arteries · Strength tests · Ultimate strain

1 Introduction

Tests connected with coronary arteries constitute a frequent subject of scientific considerations [1–12]. There has been an increase of interest in this field also in Poland in the recent years [6, 13, 14]. Nevertheless, the majority of studies are concerned with medical aspects of atherosclerosis treatment [13]. On the other hand, biomechanical research of coronary arteries is dominated by the studies relating to angioplasty and stent implantation in stenotic coronary vessels [14]. The latest cutting-edge technology and engineer's IT methods aided by modern medical diagnostics available to contemporary researchers make it possible to support the treatment of this kind of diseases by means of pre-operative planning. However, less attention has been paid to the research on the strength properties of arteries, which play an essential role in the process of creation of numerical models. Numerical simulations enable the calculation of the distribution of stress and strain in the area of the bulged wall of

B. Gzik-Zroska (✉) · K. Jozzko · W. Wolański · M. Gzik
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: Bozena.Gzik-Zroska@polsl.pl

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_26

the coronary artery. On this basis, the predictions about the probability of the pathological artery rupture can be made. It is vital especially during the designing process of stents and selection of an engineer's treatment support method. Apart from sole research purposes, the determination of the strength properties of arteries is crucial for the field of tissue engineering. The major issue connected with angioplasty planning relates to a proper selection of the balloon strain values in order to avoid the rupture of the bulged artery. Due to differences in the structure and mechanical properties, there are different guidelines for different arteries. Therefore, the more accurate knowledge of the arteries, the lower the risk connected with the medical procedure performed. Taking into account the above, this study aims to develop a testing methodology enabling the determination of mechanical properties of arteries on the example of porcine arteries using the MTS Insight 2 strength machine and digital image correlation (DIC).

2 Materials and Methods

The literature describes various methods of the determination of the mechanical properties of arteries. The first studies were published between 1880–82 [15] and were gradually developed both on a macro- and microscopic level. Young's modulus and Poisson's ratio on the macroscopic scale constitute the main data which describe the material properties of arteries. In the 20th century the most common method used for the determination of such data was a static tensile test Fung [16], Lake and Armeniades [17]. The obtained results were considerably influenced by the testing conditions [18], the distance from the heart and age Fung [16]; Tremblay et al. [19], hence the diversity of the data describing the properties of arteries. Nowadays, the tests on arteries are most often conducted in physiological conditions for a complex state of loading [20, 21]. In such conditions the measurements of strains and displacements are performed by means of contactless methods. A widely applied measuring system is a video extensometer consisting of 1 video camera [20], which makes it possible to determine the global or average data of the properties. However, a method which is used in this study on porcine artery is a 3D-DIC (Digital Image Correlation) system, which is a state-of-the-art technique [22, 23]. In this study, the assessment of the mechanical properties of coronary arteries was conducted in a uniaxial and biaxial state of loading. The DIC method was used as a measuring system during a ballooning test. The authors plan to apply the above-mentioned method to a static tensile test in the future, as it enables the determination of a three-dimensional displacement field. The testing methodology proposed makes it possible to determine the mechanical properties of coronary arteries in the conditions resembling, to a considerable degree, the angioplasty treatment.

Experimental tests were carried out on the fragments of porcine coronary arteries. In total 7 arteries were tested. They were taken from half-year-old swines in a weight group from 80 to 100 kg. Six arteries were subjected to the static tensile test: 2

branches of left anterior descending artery (LAD), 2 right coronary arteries (RCA) and 2 branches of left circumflex coronary artery (LCX). One artery, the right coronary artery (RCA) was tested with the use of digital image correlation. The sampled vessels were sterilized in an antibiotic bath and then stored at a temperature of 4 °C in plastic vials in the physiological saline solution for the maximum period of 24h until the time of the testing. Each time, the tests were conducted on the segments of arteries which were 5 cm long and did not have any lateral ostia. The specimens were subjected to a static tensile test and ballooning strength test.

The study did not analyze the influence of the samples preparation on the test results obtained.

2.1 Static Tensile Test

Examination of the arteries' strength by means of the static tensile test was carried out with the MTS Insight 2 strength testing machine (Fig. 1). It is an electromechanical machine serving the purpose of static tests, mainly low-force tests [24] in the range up to 2 kN. The specimens were subjected to a uniaxial tensile test after the ends of the specimens had been fixed in special grips. The test was performed in quasi-static conditions at a speed of 5 mm/min. Prior to testing, the initial length between the grips as well as the outer diameter of the vessels were measured. Next, the specimens were stretched along the longitudinal axis of the vessels (Fig. 1).

Response of the vessels to the load applied was recorded in diagrams, which were later analysed. During the process of stretching of the arteries segments, the values of force and displacement were recorded. An example of a tensile test chart for the

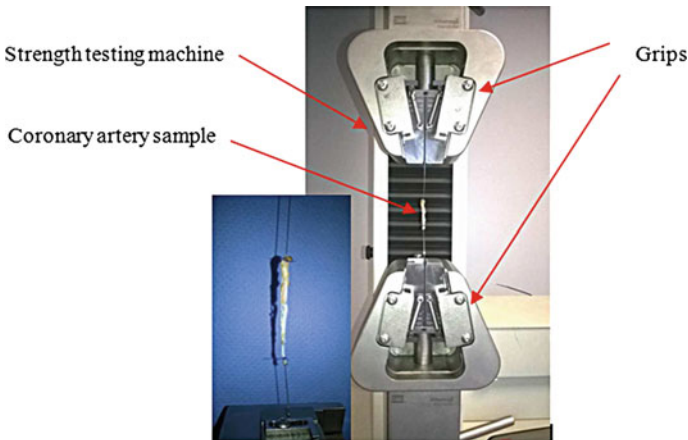


Fig. 1 Testing station

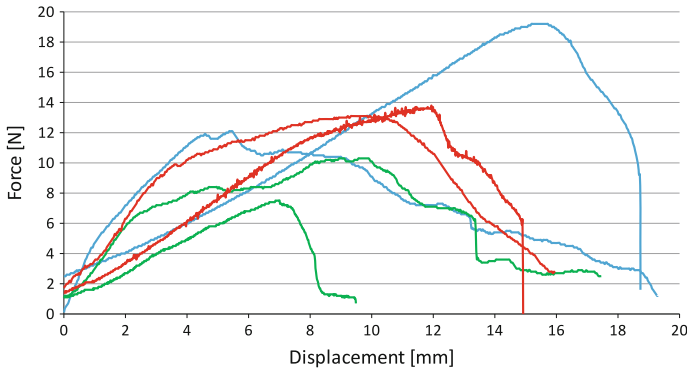


Fig. 2 An example of a force—displacement graph for LAD artery

LAD artery sample has been presented in Fig. 2. On the basis of the obtained physical values and the results analysis, it is possible to determine the general characteristics of mechanical properties of the remaining arteries.

2.2 *Ballooning Strength Tests*

Ballooning strength tests were conducted using digital image correlation on the arteries placed in the testing system including a pressure gauge (manometer). Prior to testing, an artery sample was specially prepared in order to obtain the basis for the measurement. A white layer of talc was put on the specimen's surface followed by black dots of watercolour in a chaotic arrangement. The substances used were natural and had no influence on the properties of the arteries. The Q-400 Dantec Dynamics sensor head (consisting of two cameras in a stereoscopic system, LED light source and ISTR4 4D software) [25] was used to measure strain and displacement of the specimens.

The strength tests of the arteries were conducted by inserting a small balloon inside the artery. The balloon was provided with a drain into which some liquid was pumped under increasing pressure. The level of expansion and the pressure inside the balloon were controlled by means of the manometer. Both ends of the vessel were tied to prevent slipping of the balloon out of the artery. During the test, in a controlled change of the internal pressure, the strain values were recorded as well as the displacement values of the artery walls. An example of the map of the artery strain is presented in Fig. 3.

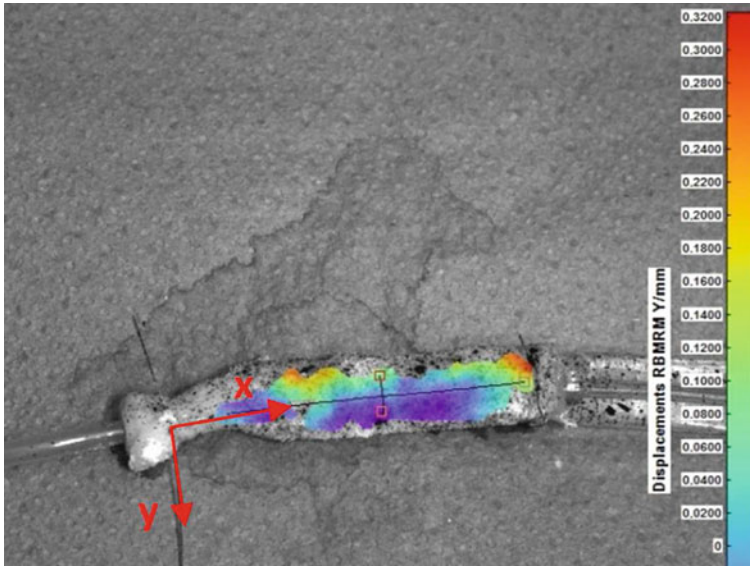


Fig. 3 Artery strain distribution map

3 Results

The experimental tests conducted made it possible to evaluate the mechanical properties of porcine coronary arteries. Figure 3 shows the results obtained in the static tensile test. For each type of the coronary artery (LCX, LAD, RCA) examined there are different maximum values of the force causing rupture. The biggest force equal to 15.6 N (± 3.55 N) was recorded for the LAD artery (Fig. 4a), whereas the LCX arteries were characterized by the biggest elongation equal to 10.63 mm (± 1.1 mm) (Fig. 4c). All types of arteries suffered destruction at a similar strain level of approximately 12.5% (Fig. 4b).

Very interesting results were obtained during the strength tests of artery ballooning with the use of digital image correlation. During the test, boundary conditions similar to the conditions of physiological load were set. Responding to the pressure change, the measurements of the strain were made in two directions x and y (Fig. 3). The strength test results obtained from the artery ballooning (Table 1) show simultaneous occurrence of transverse and longitudinal strain. The pressure change in the range of 1–9 bar results in higher values of maximum transverse strain $\varepsilon_y = 10.6\%$ ($\pm 1\%$) than maximum longitudinal strain $\varepsilon_x = 7.8\%$ ($\pm 0.6\%$).

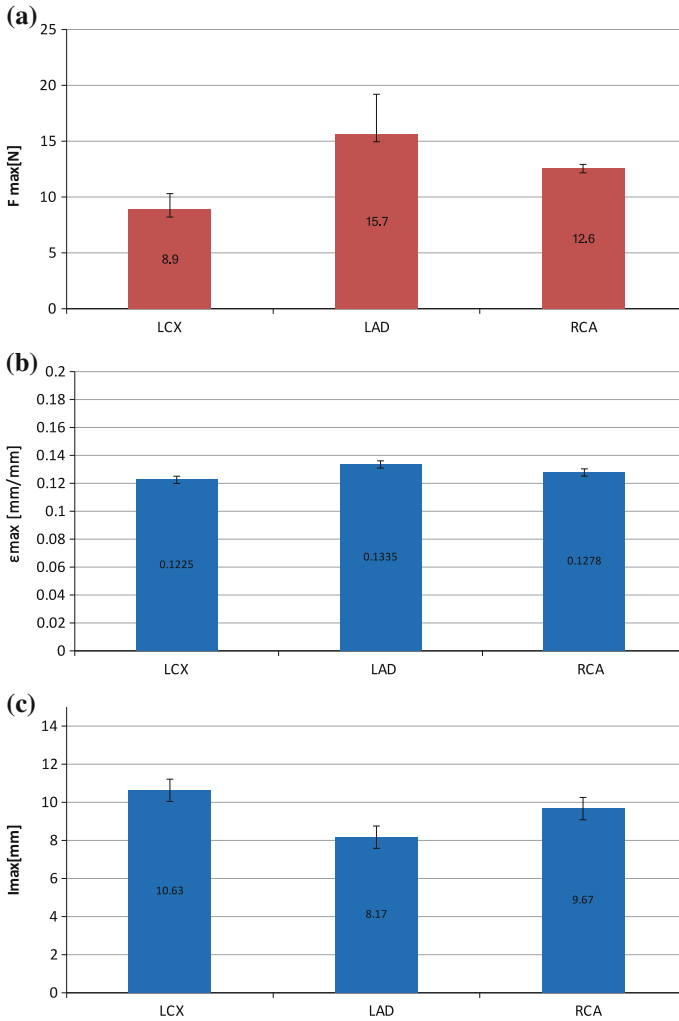


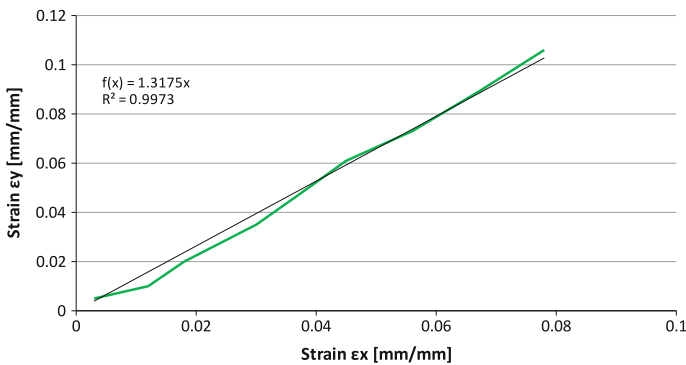
Fig. 4 Results of static tensile test performed on arteries: **a** maximum force, **b** maximum strain, **c** maximum elongation

4 Discussion and Conclusion

Understanding the mechanical properties of coronary arteries enables safe and effective treatment of their conditions. The considerations included in this work refer to the analysis of the artery strain for different kinds of load which provide information about their resistance to the load and pressure applied. There are some mentions about unidirectional strain in the literature [1, 6]. This study also presents the properties

Table 1 Results of artery ballooning strength tests

Pressure (bar)	ε_y (mm/mm)	ε_x (mm/mm)
1	0.005 ± 0.001	0.003 ± 0.005
2	0.01 ± 0.008	0.012 ± 0.005
3	0.02 ± 0.009	0.018 ± 0.005
4	0.035 ± 0.008	0.030 ± 0.005
5	0.047 ± 0.008	0.037 ± 0.005
6	0.061 ± 0.009	0.045 ± 0.005
7	0.073 ± 0.009	0.056 ± 0.005
8	0.089 ± 0.009	0.067 ± 0.006
9	0.106 ± 0.01	0.078 ± 0.006

**Fig. 5** Dependence of transverse strain ε_y on longitudinal strain ε_x

of coronary arteries obtained in the static uniaxial tensile test. It was observed that different types of arteries have a different level of resistance (maximum force) in a similar strain condition. No tests have been carried out in the scope of explaining the dependencies of the occurrence of anisotropic strains and their impact on the behaviour of an artery under load. Ballooning strength tests with the use of digital image correlation revealed that there was a linear dependence between longitudinal and transverse strains of the artery: $\varepsilon_y = 1.3175\varepsilon_x$ (Fig. 5). There is still a question to answer if the dependence observed occurs for all arteries?

Both transverse strain and longitudinal strain are the resultant resistance of the adjacent internal layers of the artery (infima, media, adventina) to the pressure applied. That is why in the strength tests performed on arteries it seems necessary to simultaneously measure both conditions of strain for the same condition of load. It is revealed by the results presented in Table 1. The knowledge of boundary values of strain allows predicting the risk of the coronary artery rupture in the planned angioplasty and stent implantation procedures. Obtaining objective information on the boundary strain values of different arteries requires further research

which would take into consideration the changes of their morphometric qualities. The research conducted by Monson [10] showed that arteries dynamically react to external factors, such as the pressure and flow of blood. In accordance with this theory, the adaptive reaction of arteries may have an impact on the strain conditions caused by the change in blood flow and pressure. The phenomena occurring in the vessels have not been fully examined yet.

Acknowledgments The study was supported by National Centre for Research and Development under grant STRATEGMED2/269760/1/NCBR/2015.

References

1. Andel, C., Pistecky, P., Borst, C.: Mechanical properties of porcine and human arteries: implications for coronary anastomotic connectors. *Ann. Thorac. Surg.* **76**(1), 58–64 (2003)
2. Carmines, V., McElhaney, J.H., Stack, R.: A piece-wise non-linear elastic stress expression of human and pig coronary arteries tested in vitro. *J. Biomech.* **24**, 899–906 (1991)
3. Gow, C., Hadfield, C.: The elasticity of canine and human coronary arteries with reference to postmortem changes. *Circ. Res.* **45**, 588–594 (1979)
4. Holzapfel, G.A., Gasser, T.C.: A new constitutive framework for arterial wall mechanics and a comparative study of material models. *J. Elast.* **61**, 1–48 (2000)
5. Holzapfel, G., Sommer, G., Gasser, C., Regitnig, P.: Determination of layer-specific mechanical properties of human coronary arteries with non-atherosclerotic intimal thickening, and related constitutive modelling. *Am. J. Physiol.-Heart Circ. Physiol.* **289**, H2048–H2058 (2005)
6. Jankowska, M., Bartkowiak-Jowska, M., Będziński, R.: Experimental and constitutive modeling approaches for a study of biomechanical properties of human coronary arteries. *J. Mech. Behav. Biomed. Mater.* **50**, 1–12 (2015)
7. Lasheras, J.C.: The biomechanics of arterial aneurysms. *Annu. Rev. Fluid Mech.* **39**, 293–319 (2007)
8. Lally, C., Reid, A., Prendergast, P.: Elastic behavior of porcine coronary artery tissue under uniaxial and equibiaxial tension. *Ann. Biomed. Eng.* **32**, 1355–1364 (2004)
9. Chow, M.-J., Zhang, Y.: Changes in the mechanical and biochemical properties of aortic tissue due to cold storage. *J. Surg. Res.* **171**, 434–442 (2011)
10. Monson, K.L., Matsumoto, M.M., Young, W.L., Manley, G.T., Hashimoto, T.: Abrupt increase in ray carotid blood flow induces rapid alteration of artery mechanical properties. *J. Mech. Behav. Biomed. Mater.* **4**(1), 9–15 (2011)
11. Pandit, A., Lu, X., Wang, C., Kassab, G.: Biaxial elastic material properties of porcine coronary media and adventitia. *Am. J. Physiol.: Heart Circ. Physiol.* **288**, H2581–H2587 (2005)
12. Sommer, G., Regitnig, P., Koltringer, L., Holzapfel, G.A.: Biaxial mechanical properties of intact and layer-dissected human carotid arteries at physiological and supra-physiological loadings. *Am. J. Physiol. Heart Circ. Physiol.* **298**, 898–912 (2010)
13. Gziut, A.I., Kulawik, T., Młotek, M.: Asymptomatic coronary fistula disclosed during angioplasty performed due to acute coronary syndrome. *Kardiol. Pol.* **66**, 105–107 (2008)
14. Kurkowska-Nouyrigat, W., Szumbaski, J.: Modelowanie przepływu krwi przez naczynia wieńcowe ze stentem (Modeling of blood flow through stented coronary arteries). *Aktualne Problemy Biomechaniki.* **3**, 115–120 (2009) (in Polish)
15. Roy, C.S.: The elastic properties of the arterial wall. *J. Physiol.* **3**(2), 125–159 (1881)
16. Fung, Y.C.B.: Elasticity of soft tissue in simple elongation. *Am. J. Physiol.* **156**, 445 (1961)
17. Lake, L.W., Armeniades, C.D.: Structure—property relations of aortic tissue. *ASAIO J.* **18**(1), 202–208 (1972)

18. Chow, M.J., Zhang, Y.: Changes in the mechanical and biochemical properties of aortic tissue due to cold storage. *J. Surg. Res.* **171**(2), 434–442 (2011)
19. Tremblay, D., Cartier, R., Mongrain, R., Leask, R.L.: Regional dependency of the vascular smooth muscle cell contribution to the mechanical properties of the pig ascending aortic tissue. *J. Biomech.* **43**, 2448–2451 (2010)
20. Ciszek, B., Cieślowski, K., Krajewski, P., Piechnik, S.K.: Critical pressure for arterial wall rupture in major human cerebral arteries. *Stroke* **44**(11), 3226–3228 (2013)
21. Holzapfel, G.A., Sommer, G., Gasser, C.T., Regitnig, P.: Determination of layer—specific mechanical properties. *Circulation* **289**, H2048–2058 (2005)
22. Badel, P., Avril, S., Lessner, S., Sutton, M.: Mechanical identification of layer-specific properties of mouse carotid arteries using 3D-DIC and a hyperelastic anisotropic constitutive model. *Comput. Methods Biomech. Biomed. Eng.* **15**, 37–48 (2012)
23. Wittek, A., Karatolios, K., Bihari, P., Schmitz-Rixenc, T., Moosdorf, R., Vogt, S., Blase, C.: In vivo determination of elastic properties of the human aorta based on 4D ultrasound data. *J. Mech. Behav. Biomed. Mater.* **27**, 167–183 (2013)
24. Kajzer, A., Kajzer, W., Gzik-Zroska, B., Wolański, W., Janicka, I., Dzielicki, J.: Experimental biomechanical assessment of plate stabilizers for treatment of pectus excavatum. *Acta Bioeng. Biomech.* **15**(3), 113–121 (2013)
25. Kokot, G., Binkowski, M., John, A., Gzik-Zroska, B.: Advanced mechanical testing methods in determining bone material. In: *Mechanika: Proceedings of 17th International Conference*, Kaunas, pp. 139–143 (2012)

Risk Factors Influencing Lower Limbs Injuries During IED Blast

Marek Gzik, Wojciech Wolański, Bożena Gzik-Zroska, Kamil Jozzko, Michał Burkacki and Sławomir Suchoń

Abstract Improvised explosive device IED attack becomes main reason of casualties among coalition soldiers during military missions in Afghanistan and Iraq. In under-belly explosion scenario the most exposed on blast effects body part is leg. The aim of studies presented in this paper has been indication of factors enhancing lower leg injury risk during IDE attack. Model of soldier in KTO Rosomak (Polish army wheeled armoured vehicle) crew compartment created in Madymo software has been used for research. Results of simulations has been analysed and injury criteria has been determined.

Keywords Soldier · Armoured vehicle · Biomechanical analysis · Injury criterion · IED blast · Madymo

1 Introduction

The lower extremity is the foremost soldier's body part injured during vehicle IED attacks [10, 17]. Even when passengers and crew survives without critical or fatal injury their lower limbs could be injured suppressing their combat abilities (Fig. 1) and prone to a high risk of debilitating injuries [6].

Shock wave generated by explosion reaches vehicle hull in approximately 0.5 ms depending on distance between explosive device and vehicle surface [7]. Beginning from bottom plating, mechanical load from shock wave is transferred on hull construction and floor plate which deform plastically and elastically (Fig. 1).

This deformations threatens occupant feet, ankle, tibia, knee, femur and pelvis due to rapidly increasing compression force. Load transferred on lower extremity reaches its maximum value in approximately 10 ms [11]. Described mechanism takes place

M. Gzik (✉) · W. Wolański · B. Gzik-Zroska · K. Jozzko · M. Burkacki · S. Suchoń
Faculty of Biomedical Engineering, Silesian University of Technology,
Roosevelta 40, Zabrze, Poland
e-mail: marek.gzik@polsl.pl
URL: <https://www.polsl.pl/Wydzialy/RIB>



Fig. 1 *Left* lower leg injury harmed in IED attack on Polish armoured wheeled transporter in Afghanistan [17]; *right* metal comb—simple device for floor displacement measurement during experiment

Table 1 Lower leg pass/fail criterion for vehicle safety evaluation [12]

Criterion	Measurement device	Pass/fail level	Signification
Tibia axial compression force	Load cell in the lower Tibia of the Hybrid III	2.6 kN	10% risk of AIS 2+ (moderate injury)

in the beginning of so called primary phase, during first 50 ms in which occupants legs are exposed on highest injury risk.

Subsequently, in primary phase global effects occur when vehicle absorbs energy from shock wave and starts moving vertically (in under-belly explosion scenario when explosive device is placed under center of hull or under carrier wheel). Depending on vehicle weight and power of explosive device effect reaches its peak in 100–300 ms. After reaching highest point vehicle starts falling down due to gravity. This is the moment when rebound phase begins. Loads caused by gravity are less significant. In some cases the third phase occurs—resulting accident which may cause events like: rollover, car crash, overturn, crash into ditch etc.

Lower leg injuries caused by: global effects, rebound phase or resulting accident often are caused by collision of leg with objects inside crew compartment. For instance, when soldiers’ seats are placed in front of each other, thrown up legs may hit seat construction [2, 3].

NATO standardization documents [11, 12] recommends as human surrogate Hybrid III 50th ATD dummy (Anthropomorphic Test Device) in mine or IED impact research. Hybrid III family is verified and commonly used in automotive crash tests. Dummy sensors fully allows determine injury criteria specified in HFM-90 and HFM-148 [11, 12] (Table 1).

Hybrid III 50th dummy is completely qualified for IED blast analysis, however H-III lower leg was created for crash test where dynamics of event and force direction is different than in blast experiment [5, 12]. H-III leg in blast experiments turn out to be too rigid compared to human leg and may overstate tibia compression force.

Mil-Lx structure includes compliant element reflecting tissues properties. According to this HFM-148 advice H-III lower leg replacement with Mil-Lx lower leg which has more biofidelic response to rapid load in vertical direction [14, 15] (Fig. 2). H-III and Mil-Lx models have been validated [5, 15].

2 Aim of Studies

The aim of studies was to identify factors influencing tibia injury risk in under belly explosion scenario of 10kg of TNT charge. Three scenarios were proposed to obtain influence and weight of following factors:

- A Soldier with feet placed on floor or footrest to obtain influence of footrest as internal safety system;
- B Soldier with floor plate deformation to determine local effects importance;
- C Soldier with opposing seat to obtain injury risk from collision with seat footrest to investigate mutual influence between passengers and interior Fig. 3.

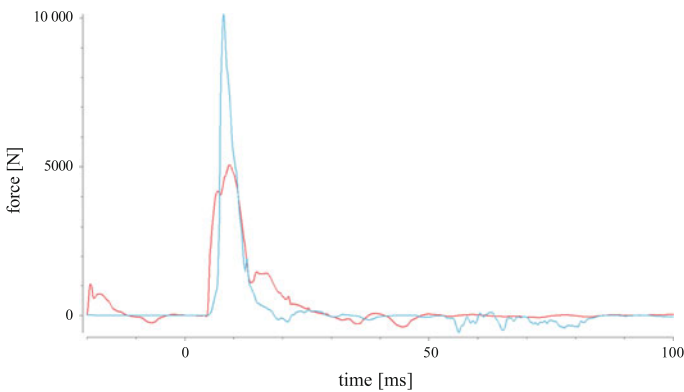


Fig. 2 Comparison of H-III lower leg (blue line) and Mil-Lx lower leg (red line) respond in numerical experiment; results coincide with literature reports [1, 12]

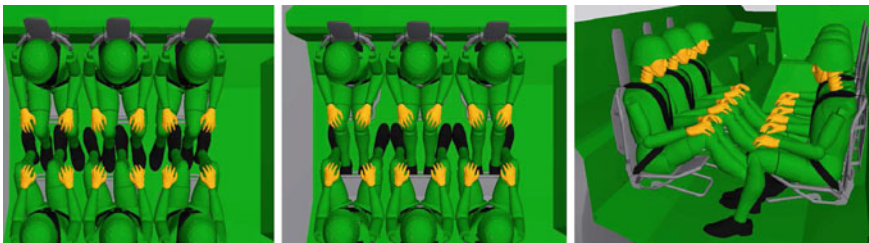


Fig. 3 Soldiers arrangement for experiment C

3 Numerical Models

Numerical models of soldier, crew compartment and vehicle has been created in Madymo[®] software for wide case analysis. However, in this paper authors present work focused on lower leg only.

Soldier model consists of Hybrid III 50th dummy model with Mil-Lx leg and wz. 2005 Polish army helmet model. Helmet model has been created basing on point cloud obtained through 3D scanning of real helmet. KTO Rosomak crew compartment and vehicle outer geometry have been modeled basing on real vehicle measures. Seats with footrest geometry have been meshed using CAD models. Safety harness have been created using Madymo[®] built-in belt fitting tool. Vehicle geometry has been made as rigid body with the exception of experiment B (Sect. 4). For experiment B floor plate has been made as FEM element. Harness, seat fabric and helmet suspension has been defined also as FEM elements with material characteristics. Footrest has been made as rigid element but with consideration of deformation characteristic (obtained in FEM simulations) which helps to reduce computational time [3]. Foot, shoe sole and floor contact characteristics corresponds with HIII dummy [5].

In model over 200 contacts with proper characteristics were defined divided into three groups:

- contacts within single dummy (e.g. dummy's A hand to dummy's A abdomen);
- contacts between dummies and environment (e.g. dummy's back to seat);
- contacts between dummies (e.g. dummy's A leg to dummy's B leg).

As input data acceleration impulse of peak value 300g effecting vehicle hull was applied to correspond with explosion effect. Floor deformation was taken from LS-Dyna[®] with Conwep algorithm simulation of 10kg TNT explosion underneath vehicle, placed 1m under soldiers feet. Mannequin kinematics and dynamics as output data allows to calculate injury risk criteria for safety evaluation.

Due to model features 200ms of simulation takes estimated 30min of computational time (regular PC with Intel i7 3rd gen. processor) which allows to simulate not only local effects and beginning of primary phase but also global effects and later events up to few seconds (like roll-over).

4 Simulations Results

In order to identify factors enhancing lower leg injury risk three variations of experiment with under-belly explosion of IED equivalent to 10kg of TNT (300g acceleration peak [13]) had been conducted (Figs. 4 and 5).

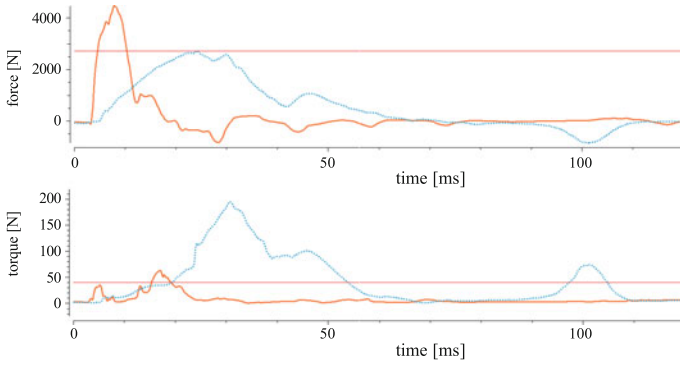


Fig. 4 Experiment A: lower tibia compression force and torque for foot placed on floor (*solid line*) and on footrest (*dotted line*), horizontal line marks 10% risk of AIS 2+ injury

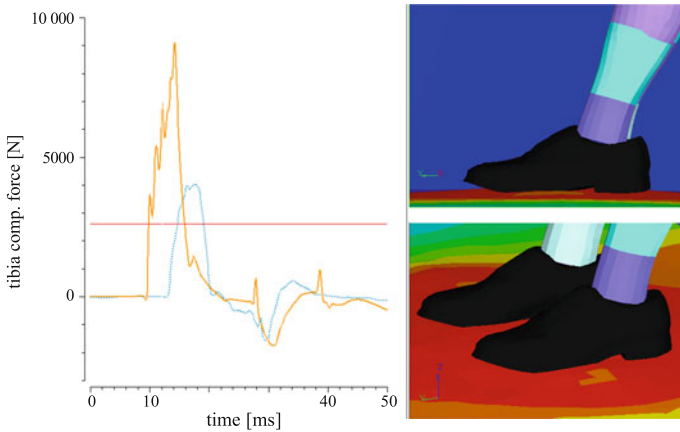


Fig. 5 Experiment B: foot placed on floor without floor deformation (*solid line*) and with floor deformation (*dotted line*)

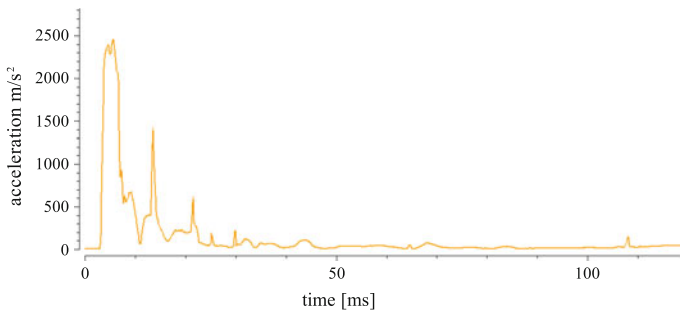


Fig. 6 Experiment C: right foot resultant acceleration; contact foot-seat starts at 80th ms

5 Conclusion

Experiment A shows that footrest significantly reduces tibia compression force from 4.48 kN to acceptable 2.49 kN level (Fig. 4). According to HFM-148 [12] this provides positive test result. In HFM-148 document [12] doesn't apply injury criteria for ankle joint however there is indicated level of 40 Nm for foot dorsiflexion (AIS 2+ 10% risk injury) [4]. In foot on floor scenario ankle joint torque reaches maximum value of 65 Nm but placing foot on footrest increases torque up to 195 Nm.

In experiment B significant difference was observed. For model with rigid floor plate tibia compression force reaches value of 4.43 kN; for floor deformation 9.11 kN (Fig. 5). Referring to Sect. 1, local effects are main threat for lower leg safety. Experiment shows that excluding local effect from model may distort results of simulation (for lower leg) and underrate tibia compression force.

In experiment C soldier's arrangement was difficult due to lack of space in crew compartment. Currently used seats and their placing doesn't allow all members to use footrest (Fig. 3). Simulation shows that legs (kept on floor and footrest) may hit seat or footrest construction from below at approximately 80th ms. Acceleration load was detected at moment of collision (Fig. 6) but there is no significant influence of this hit on injury criteria (tibia comp. force at 80th ms is 0.2 kN).

6 Discussion

From injury risk viewpoint preventing leg and floor plate contact in the moment of explosion is crucial. Footrest decrees tibia compression force for 44% (experiment A results) reducing injury risk to acceptable level. Ankle torque for dorsiflexion (subsidiary to tibia compression criterion) is increased due to footrest constitution, however precise determination of ankle injury requires soldier footwear addition as described in [11, 12]. To sum up, the most efficient way to lower risk of tibia injury is to isolate feet from floor (in case when vehicle doesn't have any energy absorbing structures).

Publication in study area describes local effects as crucial for lower leg injury mechanism [11]. Floor deformation magnifies load transfer on legs and can't be omitted in models dedicated for lower leg investigation.

In currently works many authors use FEM approach for whole model, which is computational expensive and doesn't allow to reflect full crew compartment situation. Whole body kinematic and dynamic analysis (eg. [8, 9]) may lead to safety increase. In [16] authors finds Madymo[®] software suitable for use in injury assessment. Moreover model presented in this paper allows to investigate relations between crew members and interior in complex model.

Acknowledgments The study was conducted within "Improve the safety and protection of soldiers on missions through action in military-medical and technical areas" research project co-financed by the Polish National Center of Research and Development (DOBR-BIO4/022/131449/2013).

References

1. Bailey, A., Christopher, J., Henderson, K., Brozoski, F., Salzar, R.: Comparison of Hybrid-III and PMHS response to simulated underbody blast loading conditions. In: Proceedings of IRCOBI Conference on the Biomechanics of Impact, Gothenburg, Sweden (2013)
2. Gzik, M., Wolański, W., Gzik-Zroska, B., Joszko, K., Burkacki, M., Suchoń, S.: Modelowanie żołnierza na potrzeby analizy biomechanicznej wpływu obciążenia pochodzącego od fali uderzeniowej (Soldiers model for biomechanical analysis of blast load), In Bezpieczeństwo wojsk aspekcie zagrożeń wynikających z użycia improwizowanych urządzeń wybuchowych (IED) (Troops safety in terms of IED threats), pp. 103–112, Warszawa (2014). ISBN 978-83-7523-376-6 (in Polish)
3. Gzik, M., Wolański, W., Gzik-Zroska, B., Joszko, K., Burkacki, M., Suchoń, S.: The impact of ergonomic factors influencing armoured vehicle crew safety. In: 3rd Polish Congress of Mechanics and 21st Computer Methods in Mechanics, vol. 1, pp. 223–224 (2015)
4. Kuppa, S., et al.: Lower extremity injuries and associated injury criteria. In: 17th International Technical Conference on the Enhanced Safety of Vehicles in Amsterdam, The Netherlands, National Highway Traffic Safety Administration, Washington, D.C. (2001)
5. MADYMO Human Models Manual Release 7.2, Tass International (2010)
6. Manseau, J., Keown, M.: Development of an assessment methodology for lower leg injuries resulting from antivehicular blast landmines. In: Gilchrist, M.D. (ed.) IUTAM Proceedings on Impact Biomechanics: From Fundamental Insights to Applications, pp. 33–40. Springer, The Netherlands (2005)
7. McKay, B.J.: Development of lower extremity injury criteria and biomechanical surrogate to evaluate military vehicle occupant injury during an explosive blast event, Doctoral Dissertation, Wayne State University (2010)
8. Michnik, R., Jurkojc, J., Wodarski, P., Mosler, D., Kalina, R.M.: Similarities and differences of body control during professional, externally forced fall to the side performed by men aged 24 and 65 years, Archives of Budo, vol. 10 (2014)
9. Michnik, R., Jurkojc, J., Wodarski, P., Mosler, D., Kalina, R.M.: Similarities and differences of the body control during professional collision with a vertical obstacle of men aged 24 and 65 Archives of Budo, vol. 11 (2014)
10. Owens, B.D., Kragh, J.F., Macaitis, J., Svoboda, S.J., et al.: Characterization of extremity wounds in operation Iraqi freedom and operation enduring freedom. *J. Orthop. Trauma* **21**, 254–257 (2007)
11. RTO Technical Report HFM-090, Test Methodology for Protection of Vehicle Occupants against Anti-vehicular Landmine Effects (2007)
12. RTO Technical Report HFM-148, Test Methodology for Protection of Vehicle Occupants Against Anti-vehicular Landmine and/or IED Effects (2012)
13. Sławiński, G., Niezgoda, T., Barnat, W., Wojtkowski, M.: Numerical analysis of the influence of blast wave on human body. *J. KONES Powertrain Transp.* **20**(3) (2013)
14. Technical Product Sheet MIL-Lx Leg Model 585-0000-1, Humanetics
15. User Guide Mil-Lx legform model v1.0 R7.4, Tass International (2012)
16. van der Horst, M.J., Leerdam, P.J.: Experimental and numerical analysis of occupant safety in blast mine loading under vehicles. In: Proceedings of International IRCOBI Conference on the Biomechanics of Impact. Munich, Germany, pp. 355–356. ISBN 2-9514210-3-6 (2002)
17. Wojtkowski, M., Płomiński, J., Waliński, T.: Analiza powybuchowych obrażeń narządu ruchu u żołnierzy polskich kontyngentów wojskowych—doniesienia wstępne (Analysis of after blast injuries of locomotor system among Polish troops—initial reports), In Bezpieczeństwo wojsk aspekcie zagrożeń wynikających z użycia improwizowanych urządzeń wybuchowych (IED) (Troops safety in terms of IED threats), pp. 103–112, Warszawa (2014). ISBN 978-83-7523-376-6 (in Polish)

Hybrid Method of Human Limb Joints Positioning—Hand Movement Case Study

Grzegorz Glonek and Adam Wojciechowski

Abstract Precise and unambiguous limbs motion tracking is one of the key aspects laying behind natural human-machine communication. The paper presents a novel approach to depth sensor (Microsoft Kinect) and inertial measurement units (IMU) data fusion, providing more precise and stable hand joints tracking. The new method substitutes, mainly described in literature, sensors-derived joints position fusion with sensors-derived bones orientations fusion and subsequent joints positions estimation. Obtained joints positioning precision became even 25 % better than in other solutions. The paper comprises also the method evaluation results. It was verified both against professional motion tracking VICON system and Kalkbrenner method [6], the most relevant to the presented solution.

Keywords Kinect · IMU · Fusion · Biomechanics · Joints positioning

1 Introduction

Limb motion tracking, understood as an unambiguous and delay-minimizing process of limb's joints 3D space position estimation, is a valid problem invaluable for current researches on Natural User Interface design. It is used nowadays in several areas such as entertainment (games and movies animations), interaction with scene objects in augmented reality systems or motor skills rehabilitation. This last area, supported by the computer system, constitutes a part of a wider subject named tele-rehabilitation.

For several years, the only possibility to obtain the limb joints tracking desired accuracy was to exploit professional motion capture systems i.e. VICON or Optitrack. However, since couple of years, there appear broadly available (and cheaper) devices (i.e. game controllers) that allow to track selected aspects of human motion at user's home. On the other hand, nonprofessional solutions reveal several imprecisions and

G. Glonek (✉) · A. Wojciechowski
Institute of Computer Science, Lodz University of Technology, Lodz, Poland
e-mail: grzegorz@glonek.net.pl

constraints that might be compensated by an appropriate controllers derived data fusion.

In the paper two types of such devices were taken into consideration: Microsoft Kinect 360—a RGB–D camera that is able to track whole human body and inertial measurement units (IMU) consisting of accelerometer and gyroscope sensors that are able to measure linear acceleration and angular speed. As the devices' recording frequencies are limited (30Hz for Kinect and 70Hz for IMU) and the context of rehabilitation and hand gesture controlled object manipulation is considered, the hand tracking accuracy is superior to the speed of hand movement.

Though several authors [2–4] have proved that Kinect and IMU data fusion assures limbs joints positions tracking accuracy of about 2.5–3 cm, presented method achieves better results: 2–2.5 cm.

2 State of Art

Considered sensors have several measurement characteristics that should be taken into consideration during the fusion. Microsoft Kinect controller loses its tracking ability due to body parts occlusion [1]. Moreover, while tracking, lost joints may affect the tracking accuracy of those which are fully visible to the sensor's camera. The rotation of user's body might be an example of such scenario. Basing on author's experiments, if the user rotates more than 50° (angle between the user and camera view directions), occluded shoulder joints (and almost half of the body) will be invisible to the device and measurements of visible parts will be unstable.

Another important characteristic is that joints positions measurement accuracy change with the distance between the human and the device [7]. Figure 1 shows how the estimated accuracy changes with a distance. As the most important IMU flaws, the gyroscope measurement drift and the temperature related bias in accelerometer

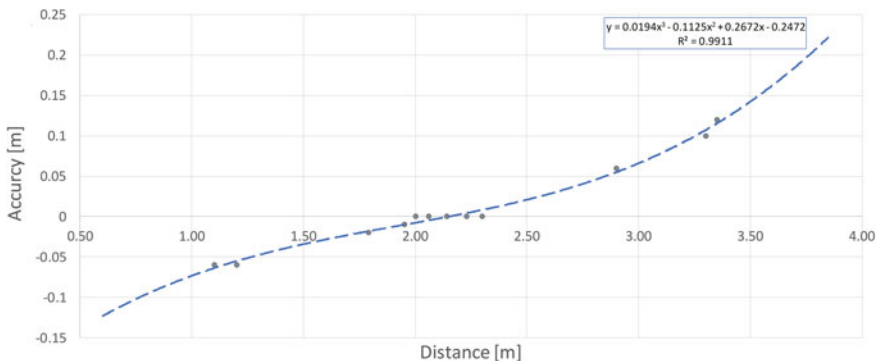


Fig. 1 Position measurement accuracy in Z-axis to distance from the Kinect

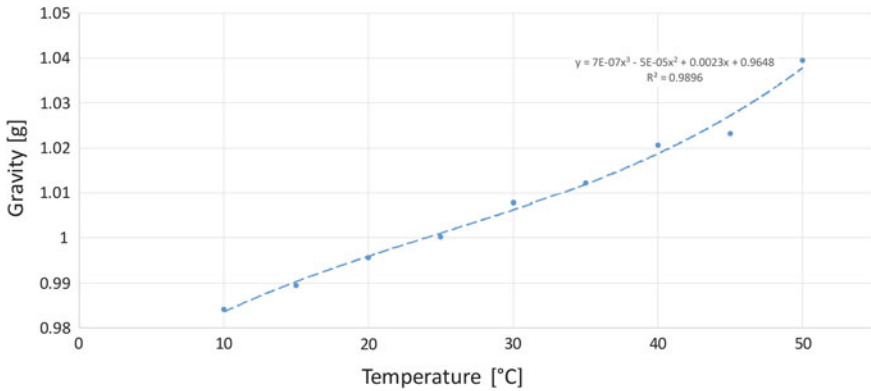


Fig. 2 Gravity measurement in temperature range 10–50°C

measurements may be indicated [11, 17]. The influence of the temperature on the bias is presented in Fig. 2.

The studied paper [9] estimates Microsoft Kinect general posture estimation variability in range 5–10 cm. Moreover, the author pointed out that the length of tracked bones vary between measurement frames in range 2–5 cm.

Considering selected controllers, several hybrid data fusion approaches, improving positioning accuracy, can be found in the literature. Authors have elaborated different approaches to sensors’ data fusion, which characterize various levels of measurements reliability. The first group of approaches can be classified as methods which use Kinect measurements as a reference system and partially relay on its measurements.

Bo et al. [2] described the joint angle (angle between joint adjacent bones) estimation method exploiting 5 degrees of freedom (DOF) IMU and Kinect. In the initial stage the method interprets the gyroscope and the accelerometer data separately and their fusion by the linear Kalman filter (KF) is subsequently performed. The same angle estimated with Kinect data is used to initially calibrate and then temporarily correct the bias of the accelerometer estimation. There were no numerical results provided in the paper, however the presented charts suggest considerable improvement of data fusion results.

A different approach was presented in a paper published by Destelle et al. [3]. Authors decided to use a set of 6DOF IMUs supported by a magnetometer, where each unit was stuck to one of the tracked limb bones. Basing on gathered measurements, orientation of each bone was estimated by the Madgwick’s algorithm [8] and their superposition resulted in the full skeleton model. Kinect data was used twofold. The first stage was to get the initial, reference skeleton frame to *label* data estimated from inertial units and to improve the IMU calibration. That process resulted in the hierarchical definition of bones orientations (*inertial skeleton*). The second stage of Kinect exploitation was to track the position of the central body point (torso joint) and then update the whole *inertial skeleton* relatively to this points displacement.

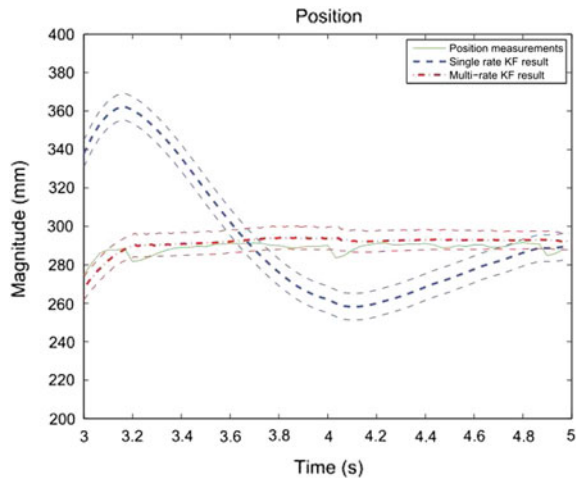
Resulting, VICON-referenced knee joints angle estimation error varied between 4° and 14° . It depended on cross correlated joints, the measurements were referenced to, as there were no joints absolute angle estimation performed.

The newest method that could be qualified to this group, is the one presented in 2015 by Tian et al. [12]. Authors included geometrical constraints of the performed motion in the estimation process, to eliminate estimations that are impossible to achieve in the real life i.e. angle between forearm and arm cannot be greater than 180° . The fusion algorithm used by authors is based on the Unscented Kalman Filter (UKF) [15]. Presented results show that the algorithm is able to work also while Kinect is outage, which was not obvious in previously described methods. To estimate method accuracy, authors compared elbow angle measurements. Authors published the information that angle measurements estimated by their fusion method deviates less than 20 degree from the expected value.

The another group of published methods is based on the assumption that both measurement devices' imperfections need to be corrected continuously by the signal fusion. In 2014 Feng and Murray-Smith [4] proposed the multi-rate Kalman filter-based fusion method of joints positions, estimated by Kinect with linear acceleration and velocity of this joint. Presented results showed that this approach stabilizes measurements around the real value much faster than single rate KF. This is visible especially when the movement starts/stops (Fig. 3). The accuracy of the presented method can be estimated around 1.5–2 cm (based on published diagrams). However, presented results refer to very short time periods (up to 5 s), so it is impossible to estimate how the method works in a long term.

A different approach was presented in the paper of Kalkbrenner et al. [6]. Authors of this publication suggested Kalman-based linear fusion of joints estimated positions retrieved from bones orientations superposed with the skeleton model (bones length)

Fig. 3 Joint position estimation with the multi-rate Kalman filter and the single rate Kalman filter [4]. Red line—multi-rate Kalman filter, blue line—single-rate Kalman filter



and Kinect measurements. Absolute joints positions estimation results were around ± 2.5 cm and seem to be the most accurate in long term experiments.

In the papers presented so far, motions, described as test movements, were performed in the *Kinect friendly* way. It means that all tracked joints were fully visible during every motion sequence and movements were done along the single axis and didn't include any occlusions.

The approach presented in 2013 by Helten et al. [5] is more advanced than methods presented so far and uses the pattern recognition to estimate the user's current pose. In the previously analyzed articles, the IMU data was always fused with Kinect skeleton data and basing on that, multiple pose features were calculated. In this article, Kinect is used as a depth camera and data from the depth stream is fused with IMU measurements. The motion recording was performed with the use of six Xsens MTx 9 DOF IMUs [18] which is a full set of professional inertial motion tracking system. The proposed method is based on the idea of the visibility model that is built from poses estimated on inertial and Kinect measurements, and then matched with predefined poses stored in the database. However, such approach seems to be useless in scenarios where you need to estimate joints positions and other limb features, as it focuses on the general pose recognition.

Our method is the most similar to the Kalkbrenner approach, but our main contribution consists in improved, weighted sensor contextual influence which results in overall better absolute joints positions estimations. The method joints absolute positioning precision of about 2–2.5 cm was verified against the ground-truth VICON system.

3 Method

A method proposed by authors, bases on the continuous linear fusion of skeleton bones orientations with respect to the current motion context. It takes into consideration controllers reliability and compensates evaluation imperfections. The proposed motion positioning method can be defined as a function f :

$$f(A, G, T, P_0^K, P_1^K, Q^K, \Delta t) \Rightarrow [p_x^F, p_y^F, p_z^F]_t, \quad (1)$$

where: A —accelerometer measurement, G —gyroscope measurement, T —temperature measurement, P_0^K, P_1^K —start and end bone joints positions measured by Kinect i.e. elbow and wrist, Q^K —bone orientation estimated by Kinect, t —current time frame, Δt —elapsed time between previous and current frame.

Orientations are contextually presented in two forms: quaternions and Euler angles, and they are transformed between these forms with respect to North and South Pole singularities. In the method authors exploited limbs joints positions ($P^K = [p_x^K, p_y^K, p_z^K]$) and bones orientations ($Q^K = [q_w^K, q_x^K, q_y^K, q_z^K]$) supplied by the Kinect device as well as accelerometer ($A = [a_x, a_y, a_z]$), gyroscope

($G = [g_x, g_y, g_z]$) and temperature (T) measurements from each IMU. Kinect joints positions and IMU based marker locations on tracked limbs are presented in Fig. 6.

In the proposed method, data gathered from measurement devices, are denoised in the first step and then used to calculate bones orientations. Then, orientations calculated from IMU devices and measured by Kinect are fused together and in the last step, bones length model is added to estimate absolute joints positions. The general overview of the orientation-based fusion process is presented in Fig. 4.

The data processing is performed in two parallel threads. The first one performs computations on the IMU data to estimate limbs orientations (quaternion) and the second retrieves Kinect skeleton bones orientations (quaternion). Their consequent, contextually-weighted and time-correlated, superposition results in fused bones quaternions values which, assuming skeleton model, can be transformed into estimated joints absolute positions. At the beginning of the first thread, the IMU accelerometer bias was corrected with the (2), due to the device operating temperature destructive influence on its measurements.

$$A' = \frac{A}{1 + \beta(T - T_0)}, \tag{2}$$

where:

- A' —corrected accelerometer measurement,
- A —accelerometer measurement,
- T —temperature measurement,
- T_0 —device reference operating temperature. For used device $T_0 = 25^\circ\text{C}$,
- β —correction factor. For used device $\beta = 0.0011$.

The value of β correction factor is the result of exploited IMU gravity regression analysis as a function of the device operating temperature (Fig. 2). Next, the corrected accelerometer data and gyroscope measurements were used to calculate quaternion of adjacent bone orientation with the Madgwick's filter [8]. The estimated orientation was then extrapolated to eliminate the observed delay. The linear extrapolation algorithm was used in this step (3).

$$[\Phi, \Theta, \Psi]' = [\Phi, \Theta, \Psi]_t + \gamma([\Phi, \Theta, \Psi]_t - [\Phi, \Theta, \Psi]_{t-1}), \tag{3}$$

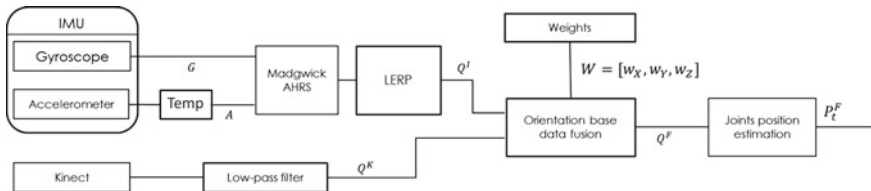


Fig. 4 Orientation based fusion method

where: $[\Phi, \Theta, \Psi,]'$ —corrected orientation in the form of Euler angles; $[\Phi, \Theta, \Psi,]$ —orientation in the form of Euler angles; γ —extrapolation factor. For used device $\gamma = 0.5$.

In the second thread, Kinect data needed to be denoised with no significant delay in measurements. It was done by first-order exponential low-pass filter defined by Eq. 4. Both joint positions and orientations have been filtered in this step.

$$y_t = \alpha x_t + (1 - \alpha)y_{t-1}, \tag{4}$$

where: y —filtered data; x —noised data; α —filtration factor. $\alpha = 0.065$.

The α factor value was estimated as a result of the analysis of the average Kinect positioning error during hand motion sequence (Figs. 5 and 6).

Both devices work in different coordination spaces (Fig. 7) and they need to be transformed into the common space before their data can be fused. As the majority of data is gathered from Kinect, its coordination space was chosen as the main one. That minimized additional computations that need to be done.

In the next step, the controllers' quaternions fusion was performed. It started with the assessment of Kinect measurements reliability. The user orientation to the camera and joints positions measurement noise level were taken into consideration. The noise level is measured by the high-pass filter in the form of Eq. (5). Sample results for keeping hands with no motion and for with lost tracking are presented in Fig. 8.

$$n_t = \delta n_{t-1} + \delta(P_t - P_{t-1}) \tag{5}$$

[16] where:

- n —noise level,
- P —joint position,
- δ —filtration factor $\delta = 0.01$.

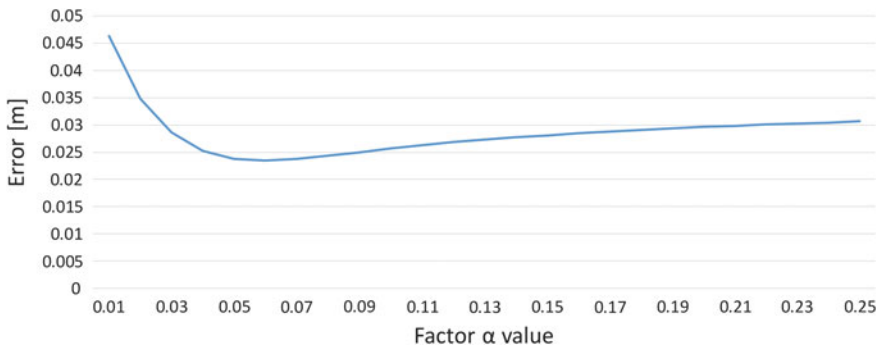


Fig. 5 Kinect measurement accuracy to low-pass filter factor α

Fig. 6 Kinect skeleton joints positions and IMU location

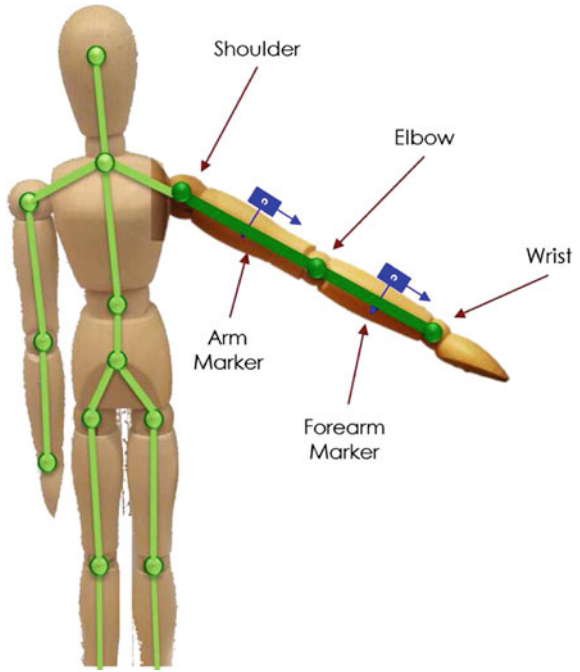
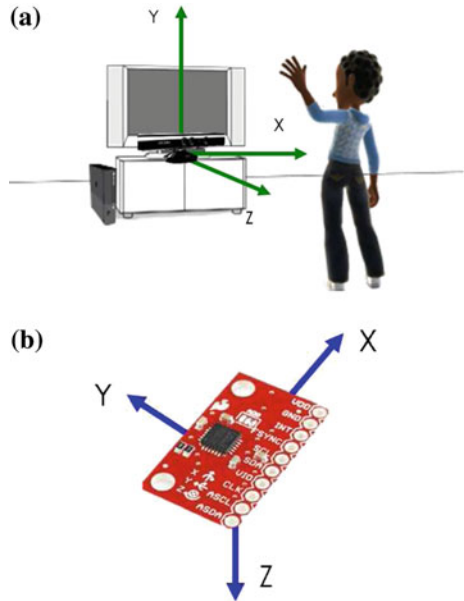


Fig. 7 Devices coordination spaces. **a** Kinect. **b** IMU



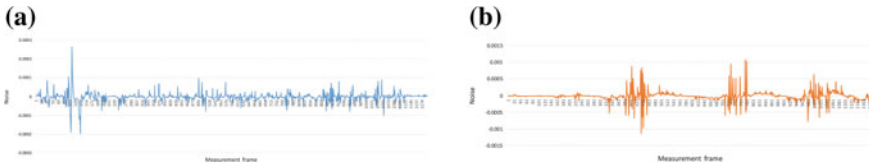


Fig. 8 High-pass filter (Eq.5) results for joint position while joint is not occluded in any time (a) and partially occluded (b). **a** Not occluded joint. **b** Occluded joint

If the user is rotated to Kinect camera more than 50° (the angle between the line of user’s shoulders and the camera surface) or the noise level is too high ($|n| > 0.0004$ based on performed experiments) then Kinect measurements are classified as unreliable and are replaced with the difference between current and previous IMU-based orientation estimations. The orientations fusion is defined as the complementary process where rotations around each axis are joined with different weights. This approach was motivated by the fact that the controllers have different measurement abilities and accuracy along each axis.

If current Kinect’s data is classified as reliable, the fusion is expressed by the following Eq. (6).

$$\begin{bmatrix} \Phi^F \\ \Theta^F \\ \Psi^F \end{bmatrix}_t = \begin{bmatrix} w_x & 0 & 0 \\ 0 & w_y & 0 \\ 0 & 0 & w_z \end{bmatrix} \begin{bmatrix} \Phi^I \\ \Theta^I \\ \Psi^I \end{bmatrix}_t + \begin{bmatrix} 1 - w_x & 0 & 0 \\ 0 & 1 - w_y & 0 \\ 0 & 0 & 1 - w_z \end{bmatrix} \begin{bmatrix} \Phi^K \\ \Theta^K \\ \Psi^K \end{bmatrix}_t \quad (6)$$

where: $[\Phi^X, \Theta^X, \Psi^X]^T$ —Euler form-based orientation; X denotes: F —fused, I —IMU, K —Kinect, and w_x, w_y, w_z : weights defining fusion factor of each axis rotation. Defined as $[0.98, 0.05, 0.65]$ respectively.

Weights used in Eq. 6 describe the level of importance of IMU measurements and need to be < 1 . The higher value used, the more important measurement was. As Kinect is not able to measure rotation along ‘x’ axis (Roll), weight close to 1 has been used there. In case of usage of inertial sensors without magnetic sensor support, rotation around gravity vector (‘y’ axis—Yaw) contains uncompensated time-related drift, so, in this case, IMU measurement was discriminated. Third axis rotation was measured by both devices, however IMU had slightly better accuracy than Kinect which was reflected in weight > 0.5 .

As both devices, Kinect and IMU, work with different sampling frequency, decimation technique has been used to pick the measurements from the closest time frames.

In case of Kinect data unreliability, the fusion was made between previously fused value and the IMU-based orientation update—between the previous and the current measurements. The fusion formula is defined as follows (7):

$$\begin{bmatrix} \Phi^F \\ \Theta^F \\ \Psi^F \end{bmatrix}_t = \begin{bmatrix} \Phi^F \\ \Theta^F \\ \Psi^F \end{bmatrix}_{t-1} + \begin{bmatrix} w_x & 0 & 0 \\ 0 & w_y & 0 \\ 0 & 0 & w_z \end{bmatrix} \left(\begin{bmatrix} \Phi^I \\ \Theta^I \\ \Psi^I \end{bmatrix}_t - \begin{bmatrix} \Phi^I \\ \Theta^I \\ \Psi^I \end{bmatrix}_{t-1} \right) \quad (7)$$

In this case w_x value remains the same and it equals 0.98 and w_y and w_z get low over the time according to the following function (Eq. 8):

$$w = \left(1 - \frac{\Delta t_n}{10} \right) \cdot 0.65, \quad (8)$$

where w —current weight value, Δt_n —amount of time in seconds when Kinect stays unreliable.

When the controllers-fused orientation is estimated it must be recalculated to the quaternion form. Then, basing on the known bone length, the position of the desired joint might be calculated.

4 Results

In order to verify the elaborated method (orientation-based joints position estimation) precision several experiments were performed. They were conducted with the VICON motion capture system as a source of a ground-truth reference data. Five users were monitored with Kinect controller, two IMUs attached to arm and forearm bones (Fig. 6) and passive marker-based VICON system, simultaneously. Markers were attached to the hand according to the schema presented in Fig. 9.

The PC used to record Kinect and IMU data was a 2.5 GHz Intel Core i7-4710HQ CPU base computer with 8GB of RAM and SSD drive. The exploited Kinect device was a dedicated Xbox 360 console controller. The software was implemented on .Net Framework 4.5 with Kinect SDK v. 1.8. IMUs—these were MPU6050 devices set up with scale ranges: ± 4 g for accelerometer and $\pm 500^\circ/s$ for gyroscope. Inertial

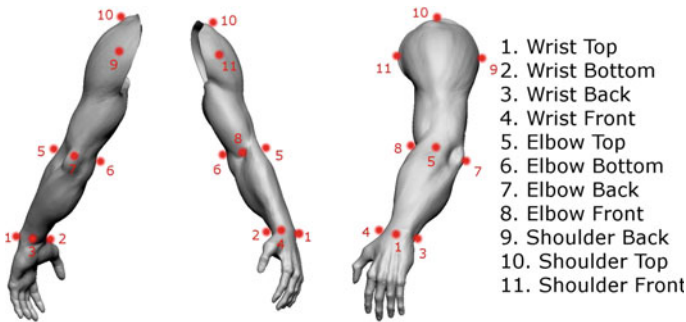


Fig. 9 Used VICON Arm marker model (Pai [10])

devices worked as a part of the self-made measurement device, built on the Arduino Due platform. The data transmission between IMUs and the PC was done through Bluetooth v. 2.0.

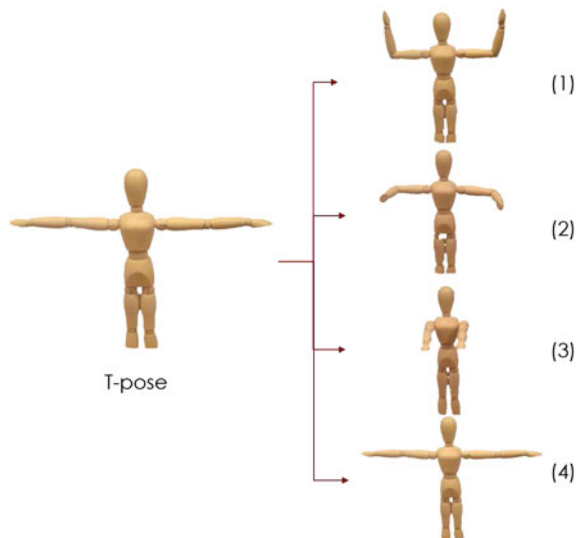
As a reference, measurements obtained from the optical multi-camera VICON motion capture has been used. Such systems are broadly used in the industry as well as to track the motion for i.e. biomechanical research [14]. According to data sheet published by the producer, declared precision of this system is down to 0.5 mm of translation and 0.5° of rotation with refresh rate up to 250 fps [13]. Such accuracy allows to use such measurements as ground-truth data. In performed experiments, VICON system worked with refresh rate 100 fps.

Performed experiments examined the right hand joints (elbow, wrist) positioning precision and the angle between the arm and the forearm estimation (the angle in the elbow joint) during 4 different movement sequences (Fig. 10):

1. Elbow flexion up to an angle of 90° .
2. Elbow flexion forward to an angle of 90° .
3. Straighten the hand in front of a body.
4. Stand still for more than a minute.

Selected gestures comprised motions that might be considered as challenging especially for Kinect. Each movement sequence started from the initial T-pose and was performed multiple times and averaged. The proposed method was also compared with the Kalkbrenner method precision, which was implemented according to the description included in the article [6]. The position estimation accuracy for elbow and wrist joints as well as the angle measurement accuracy has been presented in Figs. 11, 12 and 13. The gray color was used for Kalkbrenner position-based method results and the orange for the author's, orientation-based method achievements.

Fig. 10 Movement sequences performed during tests



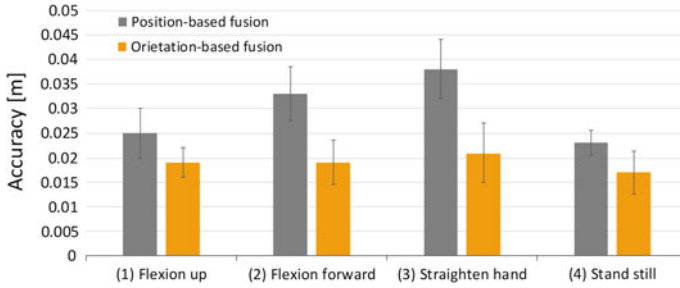


Fig. 11 Elbow positioning average accuracy

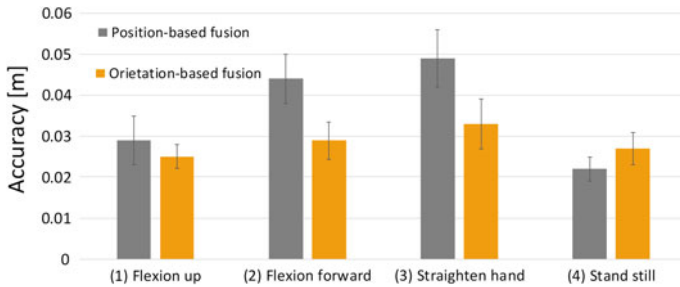


Fig. 12 Wrist positioning average accuracy

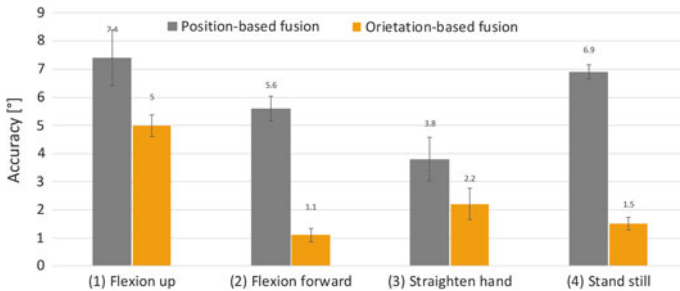


Fig. 13 Elbow angle measurement average accuracy

Unfortunately, during performed tests, authors were not able to achieve the described accuracy for the Kalkbrenner position-based method, however, the achieved results were close to the declared ones. Results presented on charts show the improvement in both: the position and the angle measurement accuracy. The average accuracy for the position estimation was 2.5 cm for the elbow and 2.9 cm for the wrist. The average angle measurement accuracy was 2.5°. The same values for position-based fusion estimations were 2.8 cm, 3.6 cm and 5.9° respectively.

5 Conclusion

The authors presented a new, orientation-based method for skeleton joints positioning. It was tested on variety of right hand movements. Obtained results have proven that the method managed to compensate imperfections of both measurement devices much better than previous approaches. Basing on the comparison of results gathered from the orientation-based fusion and the position-based fusion, the improvement of the estimation accuracy has been noticed and reached the rate of 15 up to 25 %.

Obtained results prove that the novel data fusion approach, based on the bones orientation, might be considered as an improved alternative to the well known, joint position-based methods.

References

1. Asteriadis, S. et al.: Estimating human motion from multiple Kinect sensors. In: 6th International Conference on CV/Computer Graphics Collaboration Techniques and Applications—MIRAGE '13. ACM, pp. 3–8 (2013)
2. Bo, A.P.L. et al.: Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3479–3483 (2011)
3. Destelle, F. et al.: Low-cost accurate skeleton tracking based on fusion of Kinect and wearable inertial sensors. In: EUSIPCO, pp. 371–375 (2014)
4. Feng, S., Murray-Smith, R.: Fusing kinect sensor and inertial sensors with multi-rate Kalman filter. In: IET Conference on Data Fusion and Target Tracking 2014: Algorithms and Applications, pp. 1–8 (2014)
5. Helten, T. et al.: Real-time body tracking with one depth camera and inertial sensors. In: 2013 IEEE International Conference on CV, pp. 1105–1112 (2013)
6. Kalkbrenner, Ch. et al.: Motion capturing with inertial measurement units and kinect—tracking of limb movement using optical and orientation information. In: Proceedings of the International Conference on Biomedical Electronics and Devices, pp. 120–126 (2014)
7. Khoshelham, K., Elberink, S.O.: Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors* **12**, 1437–1454 (2012)
8. Madgwick, S.O.H.: An efficient orientation filter for inertial and inertial/magnetic sensor arrays. Rep. x-io and University of Bristol (UK) (2010)
9. Obdrzalek, S. et al.: Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1188–1193 (2012)
10. Pranay, P.: Human Male Sculpt (Maya and Zbrush Pipeline). <http://graphy-x-medley.blogspot.com/2013/07/human-male-sculpt.html>. Accessed 01 Jan 2016
11. Renaut, F., Nikolic, J.: MEMS Inertial Sensors Technology. Swiss Federal Inst. of Tech. Zurich, Autonomus Sys. Lab, Studies on Mechatronics (2013)
12. Tian, Y., et al.: Upper limb motion tracking with the integration of IMU and Kinect. *Neuro-computing* **159**, 207–218 (2015)
13. Vicon: Bonita Motion Capture Camera/VICON. <http://www.vicon.com/products/camera-systems/bonita>. Accessed 12 Feb 2016
14. Vicon: Case Studies/VICON. <http://www.vicon.com/case-studies>. Accessed 12 Feb 2016
15. Wan, E.A., van der Merwe, R.: The unscented Kalman filter for nonlinear estimation. In: Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium, pp. 153–158 (2000)

16. Wikipedia Contributors: High-pass filter—Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/High-pass_filter#Algorithmic_implementation. Accessed 01 Jan 2016
17. Woodman, O.J.: An Introduction to Inertial Navigation University of Cambridge, Computer Lab Technical Report, vol. 696, pp. 1–37 (2007)
18. XSense: MTx—Products—Xsens 3D motion tracking. <https://www.xsens.com/products/mtx/>. Accessed 10 Nov 2015

A Survey of Selected Machine Learning Methods for the Segmentation of Raw Motion Capture Data into Functional Body Mesh

Magdalena Pawlyta and Przemysław Skurowski

Abstract The paper describes the progress in the research for the automatic inferring method of the body structure—functional body mesh. In the paper we investigate four motion measures and machine learning methods—variants of Gaussian mixture models, DBScan and Neural Networks. The results were analyzed both—quantitatively and qualitatively using complete and incomplete data, of healthy and impaired persons. All the learning methods were on par with the others, however, we identified cases for which certain method works better.

Keywords Machine learning · Motion analysis · Body segmentation

1 Introduction

The motion capture technology (Mocap) plays important role in the biomechanical analysis of human performance for the medical and sport applications [11, 12]. The up-to-date prominent technology is the optical motion capture [6, 7] which acquires trajectories of tracked points—markers. These markers are then organized into predefined relevant structure based on the underlying skeleton of actor, where markers located on the common body parts are within the same mesh segments and body parts are interconnected into tree hierarchy. Next, the skeletal structure is suited and driven by the appropriate marker locations. This pipeline (Fig. 1) has a drawback, any modification to the marker number or location, not to mention different species or human with strong deficiencies, needs manual editing of the representative mesh, which can be time consuming.

M. Pawlyta (✉)

Polish-Japanese Academy of Information Technology, Bytom, Poland

e-mail: mpawlyta@pjatk.edu.pl

URL: <http://bytom.pja.edu.pl/>

P. Skurowski

Faculty of Automatic Control, Electronics and Computer Science,

Institute of Informatics, Silesian University of Technology, Gliwice, Poland

URL: <http://inf.polsl.pl/>

© Springer International Publishing Switzerland 2016

E. Piętka et al. (eds.), *Information Technologies in Medicine*,

Advances in Intelligent Systems and Computing 472,

DOI 10.1007/978-3-319-39904-1_29

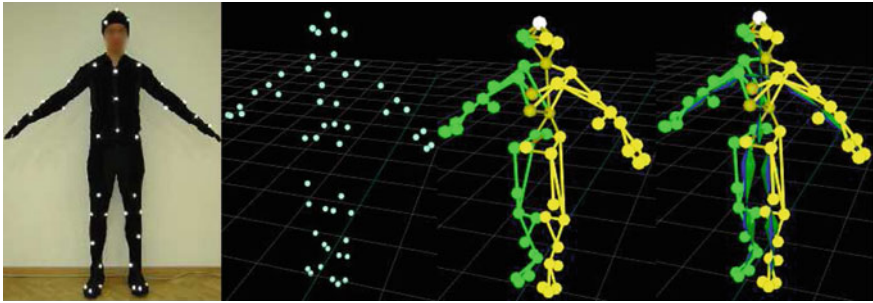


Fig. 1 Mocap pipeline: an actor, raw markers, arbitrary mesh, skeleton assigned within

Authors' former works on *functional body mesh* (FBM) [8, 9] propose a hierarchical mesh structure (Fig. 2a) which represent kinematic structure of a body which can be obtained in semi-automated way. Such a mesh forms a framework for further mocap data processing as it has advantages of both conventional representations—marker-wise—keeping all the registered data and skeletal—knowledge of the body structure. In the paper we demonstrate the progress we achieved in fully automatic inferring of the FBM.

The survey covers—brief introduction to the already existing FBM inferring method, presentation of rationales for the motion measures we use, dimensionality reduction, the results we obtained for the: linear discriminant analysis (LDA), DBScan, local and global Gaussian mixture models (GMM), artificial neural networks (ANN).

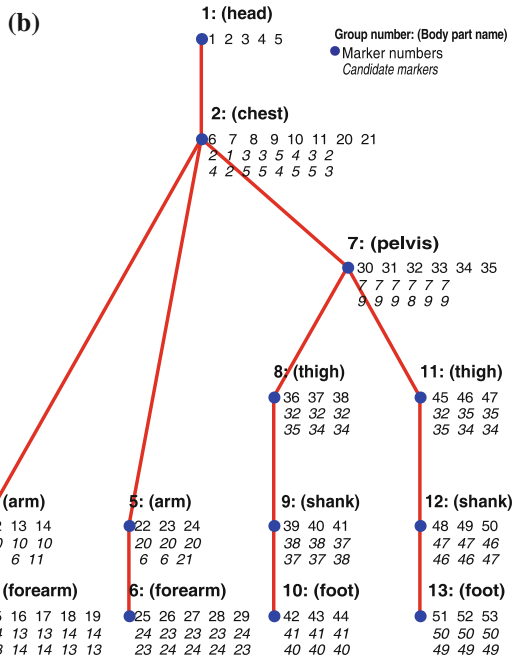
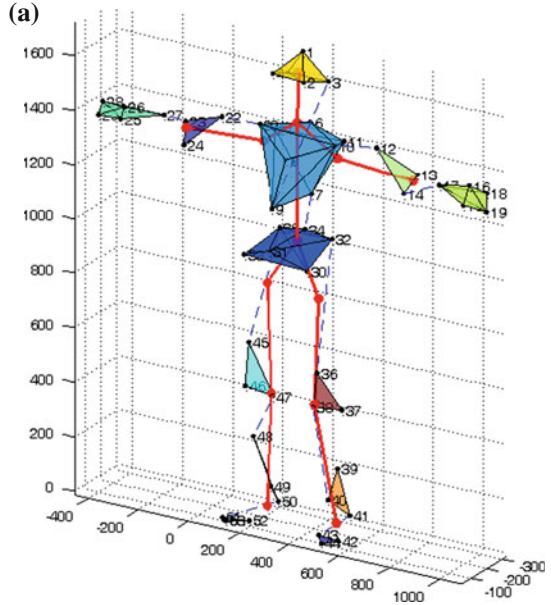
2 Method

2.1 Background

The genuine proposal is not based on any prior assumptions on the subject but the rigid body model for the inferring method. FBM forms as a hierarchy representing tree, containing groups of rigid interconnected markers, describing location of certain body parts. This approach is very flexible as it can adopt to any vertebrate subject.

Outline of the solution is as follows—in order to achieve meaningful hierarchy in the first step the markers are sorted according to the body principal axis in top-down and center-boundary order. Then are identified. The segmentation is based on two assumptions. Siblings are located on common body parts, so they move together therefore their gradients are similar, and they are located on rigid parts so their relative distance is constant. The clustering is to join markers iteratively with the thresholds on the basis of GMM model for each of the criteria. Finally, the hierarchy of the tree is recovered with a consensus voting (Fig. 2b) of group members, where each sibling

Fig. 2 Functional body mesh: **a** visualization with corresponding skeleton; **b** marker hierarchy (with children-parent voting)



group has a single parent that: is located in another group and cannot be located in a child node (no loops), The representative marker is the closest and constant distant single point to the proposing one and the siblings and parent body parts are connected so distances cannot vary very much.

2.2 New Proposal

In the paper we propose to improve obtaining method of the FBM, in order to achieve fully automatic process. We propose to extend the basic model with two additional measures—motion correlation which exhibits rotation and muscular contractions and average spatial proximity which would differentiate markers that move in-phase but are distant. Then to use PCA to reduce dimensionality, next normalize the results and employ one of the analyzed methods for a single dimensional problem. Finally aggregate the results using consensus heuristics.

3 Measuring Intermarker Relationships

3.1 Measures

Constancy of Distances—for the measurement of a distance we used Euclidean distance. Constant distance (Fig. 3a) over the whole sequence was verified for each pair of markers statistically as a range of values. To filter out noise existing in records we used the inter-quantile distance between Lower ($L = 0.5$ th) and upper ($U = 99.5$ th) percentile (P) of the whole N frame sequence:

$$D_{A,B}^E(n) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}, \tag{1a}$$

$$R_{AB}^E = P_U(D_{A,B}^E(1, \dots, N)) - P_L(D_{A,B}^E(1, \dots, N)), \tag{1b}$$

where: $x_A, y_A, z_A, x_B, y_B, z_B$ —coordinates of respective points A, B for n th frame.

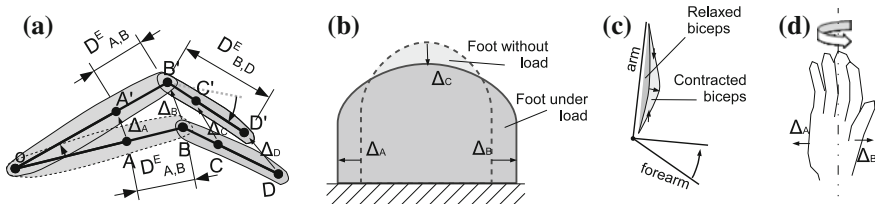


Fig. 3 Measures' rationales: **a** rigid body, **b** body plasticity, **c** muscular contraction, **d** rotation

Angular Coherence of Gradients—the gradient of the k th marker is described as:

$$\Delta_k(n) = [\Delta_x, \Delta_y, \Delta_z] = [x_n - x_{n-1}, y_n - y_{n-1}, z_n - z_{n-1}], \tag{2}$$

where: $n, n - 1$ —number of successive frames; x_n, x_{n-1}, \dots —marker coordinates in two successive frames; $\Delta_x, \Delta_y, \Delta_z$ —differences of respective coordinates in these frames, where L_k is distance of a movement of k marker is $L_k(n) = |\Delta_k(n)|$.

Cosine distance measures vectors conformance. For an ideal rigid body, pure cosine distance would be satisfactory, but during tests, it appeared necessary to use weighting by length of gradient. The cosine distance the length of the movement agnostic, meanwhile markers placed on the same body part could have small non-consistent (opponent) movements due to deformation (Fig. 3b) caused by elasticity of the human body. To suppress the influence of these deformations we decided to use weighted cosine distance where as the value of weighting function w we used averaged length of gradients. To suppress noise we employed again the interquartile range.

$$D_{\Delta_A, \Delta_B}^c(n) = w(1 - \cos(\Delta_A \Delta_B)) = w \left(1 - \frac{(\Delta_{Ax} \Delta_{Bx} + \Delta_{Ay} \Delta_{By} + \Delta_{Az} \Delta_{Bz})}{\sqrt{\Delta_{Ax}^2 + \Delta_{Ay}^2 + \Delta_{Az}^2} \sqrt{\Delta_{Bx}^2 + \Delta_{By}^2 + \Delta_{Bz}^2}} \right) \tag{3a}$$

$$w = w(\Delta_A, \Delta_B, n) = \frac{1}{2} (L_A(n) + L_B(n)), \tag{3b}$$

$$R_{AB}^c = P_U(D_{A,B}^c(1, \dots, N)) - P_L(D_{A,B}^c(1, \dots, N)). \tag{3c}$$

Correlation of Motion addresses two problems in motion coherence—rotational opponent directions of gradients (Fig. 3d) and changes in body shape (Fig. 3b, c)—such as muscular contradiction-relaxation and feet flattening under load.

$$D_{A,B}^P = 1 - \max\{r(\Delta_{Ax}, \Delta_{Bx}), r(\Delta_{Ay}, \Delta_{By}), r(\Delta_{Az}, \Delta_{Bz})\}, \tag{4a}$$

$$r(\Delta_{x_1}, \Delta_{x_2}) = \frac{\text{cov}(\Delta_{x_1}, \Delta_{x_2})}{\sqrt{\text{var}(\Delta_{x_1})} \sqrt{\text{var}(\Delta_{x_2})}} \tag{4b}$$

Proximity is another measure which is based on the fact that body parts are rather compact entities—this measure was included as an add-on as during preliminary tests, it appeared that resulting groups are too cluttered. We employed simple mean inter marker distance in whole sequence:

$$D_{A,B}^M = \frac{1}{N} \sum_{i=1}^N D_{A,B}^E(i) \tag{5}$$

All above measures were represented in a distance matrix form (Fig. 4) where markers of common body parts are located neighboring o each-other. All of the can be used as a components o vector describing the relationship of a pair of markers. In further tests we employed two scenarios—matrix D_4 comprising all the distances and

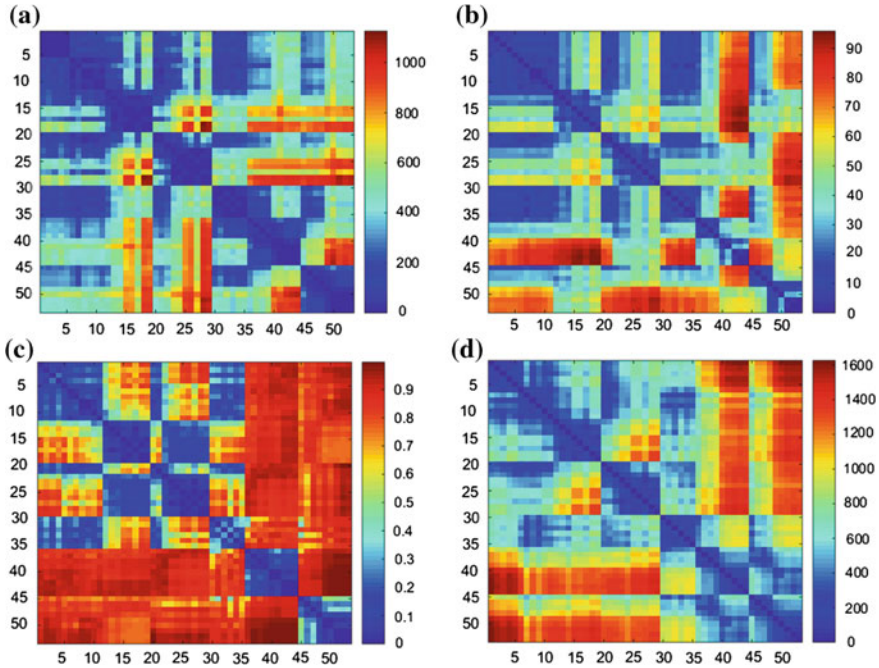


Fig. 4 Measures in a form of distance matrices: **a** range of distances, **b** range of gradient weighted cosine, **c** max gradient correlation, **d** mean distance

D_3 with proximity excluded. These matrices have four or three columns respectively, the number of rows depends on the number of markers (k) as it is a number of distances to all other markers ($\frac{k^2-k}{2}$).

3.2 Linear Discriminability

One of the first and canonical approaches was to employ the Fisher’s linear discriminant analysis (LDA) [1]—its optimal base intended for the maximum discriminability of objects where new dimensions (LD_i) are sorted according to their ability to discriminate classes (eigenvalues). The problem is generalized eigenvector decomposition:

$$A = V \Lambda V^T, \tag{6}$$

where: Λ is eigenvalues matrix, V is matrix of eigenvectors of A defined as $A = \Sigma_w^{-1} \Sigma_b$ —product of inverse of intra-class variance and between-class variance:

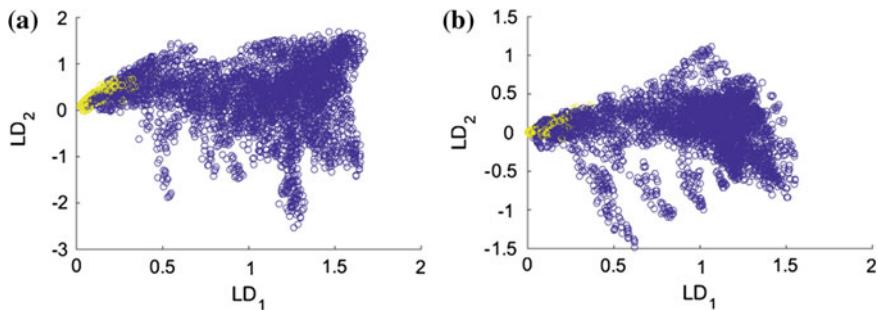


Fig. 5 Linear discriminability using: **a** D_4 , **b** D_3

$$\Sigma_b = \frac{1}{C} \sum_{i=1}^C (\mu_i - \mu)(\mu_i - \mu)^T. \quad (7a)$$

$$\Sigma_w = \sum_{i=1}^C \Sigma_i, \quad \text{where } \Sigma_i = \frac{1}{N_i} \sum_{x \in C_i}^{N_i} (x - \mu_i)(x - \mu_i)^T. \quad (7b)$$

It appeared that regardless we use D_3 or D_4 the problem is poorly separable in the linear way (Fig. 5), the results are so entangled that it is hard to identify the well separating line or surface in LDA domain.

3.3 Dimensionality Reduction

Multidimensional data of a nonuniform types are inconvenient for the most machine learning methods. When it is hard to compare samples with any distance function of Minkowski class—with special emphasis on Euclidean one. In such cases one might employ normalized (with variance) Euclidean or even Mahalanobis which employs full covariance matrix, but the most comprehensive approach is to reduce the dimensionality of the problem with respect to the data ranges and covariances with use of principal component analysis (PCA) [1] and solve it as one dimensional problem. The PCA does its task by diagonalizing the covariance matrix (Σ) into a matrix (Λ) of eigenvalues (λ_i) and produces the new orthogonal base vectors (PC—eigenvectors) which points out the direction of data variability stored as columns in change of basis matrix (V).

$$\Sigma = V \Lambda V^T. \quad (8)$$

Eigenvalues and corresponding eigenvectors (PC) are sorted in descending order, so the first principal component (PC_1) contains the most of overall variance (λ_1)

In order to control the range of input data and to fulfill conditions of a distance function it is necessary to normalize the results of PCA. The data must conform the ground (zero) level and the distance values must be positive—it is not necessary fulfilled by the results obtained from the numeric procedure. First, to obtain zero value we recalculate zero vector to the PC domain (y_0), then we normalize the values to the maximal value with element by element division of a vector:

$$y' = \left| \frac{y - y_0}{y_{\max}} \right|. \quad (9)$$

3.4 Selected Machine Learning Methods

In order to determine the relationships between the markers used three different methods representing both supervised and unsupervised approaches. As an input we used full D_3 or D_4 in NN or their PC_1 for other methods.

GMM Models Inter-marker relationship can determined by assigning of their pairs into classes according to the rationales mentioned in Sect. 3.1. One can intuitively assign each pair into one of C —four different classes (c_1, \dots, c_4): peer, close, independent and opponent. The distance functions reflect these classes so one should consider a multimodal probability distribution function (PDF) for both distance functions. The Gaussian mixture models [5] appeared to fit the data very well as it is demonstrated in [8]. The GMM is a mixture model summing two or more Gaussian probability distributions (g_i):

$$G = \sum_{i=1}^C g_i = \sum_{i=1}^C w_i N(\mu_i, \sigma_i), \quad (10)$$

where: C number of components (in our case c_1, \dots, c_4), w_i —weight, $N(\mu_i, \sigma_i)$ normal probability distribution of a μ_i , mean value and σ_i std deviation. These parameters are estimated in the process of *expectation maximization* (EM) [4]. In order to address the problem of EM initialization [10], we predefined the ‘good’ initial value with multistart using part of the data. Classification using GMM is a selection of a class with the highest mode value for the given object (x):

$$c(x) = \operatorname{argmax}_{i=1, \dots, C} \{w_i N(\mu_i, \sigma_i)(x)\}. \quad (11)$$

We employed the GMM in two ways—global and local approach. In the global approach the GMM was estimated for all the distances between markers, whereas in the local approach the GMMs are estimated separately for the distances from the current (base) marker to all other markers. The Fig. 6 demonstrates classification estimated in two different ways—global and local for a single marker.

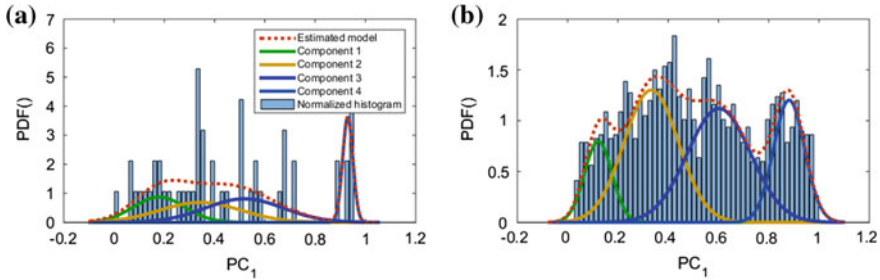


Fig. 6 GMM model using D_4-PC_1 **a** local for marker ARIE (*top of head*), **b** global

The local approach in some cases was incorrect—for some markers there was more or less classes (Fig. 6a). Using this approach, required to use of additional step to determine for each marker proper relations class. Creating such a model would take a much more time, so we decided to neglect this approach. The global approach is analogous to the genuine algorithm without tuning parameter (Fig. 6b). For the global GMM we observed the best results for the D_4-PC_1 input data (Fig. 8b).

DBScan-FC Selection of First Cluster of the Density-Based Spatial Clustering of Applications with Noise [3] which is a density based clustering algorithm. DBSCAN requires two parameters: *Eps* and the minimum number of points required to form a cluster (*minPts*). The points form a cluster if there are *minPts* or more other points within *Eps* range. In case *minPts* is 2—can't exist a group with only one marker and single outstanding points are interpreted as noise. *Eps* can be estimated using:

$$Eps = \left(\frac{\left(\prod_{i=1}^{\max(X)-\min(X)} i \right) \cdot minPts \cdot \gamma(0.5 \cdot n + 1)}{m \cdot \sqrt[n]{\pi^n}} \right)^{\frac{1}{n}}, \quad (12)$$

where: X is the m by n input matrix and γ is gamma function.

We employed DBScan-FC in a local (per marker) way—similar to the local GMM. The clustering was performed as the one versus all—for each of markers we performed DBscan on the distances to all other markers (single row in distance matrix). As a result we obtained a set of groups of which the first—closest one was the significant result. The base point due to zero distance is always identified as a noise. For the DBScan-FC the best results were observed using D_3-PC_1 input data (Fig. 8c).

Feed Forward Neural Network In order to test the fully supervised and nonlinear approach we employed feed forward neural network (FFNN), it is the simplest type of artificial neural networks where connections between units don't form a cycle or loops (Fig. 7). Information moves only in one direction—forward—from input node through the hidden nodes and to the output nodes. As an input we provided all the four distances for each of the marker pairs. The number of hidden layers and the number of neurons on each layer (n, m) is crucial [2] and was determined empirically

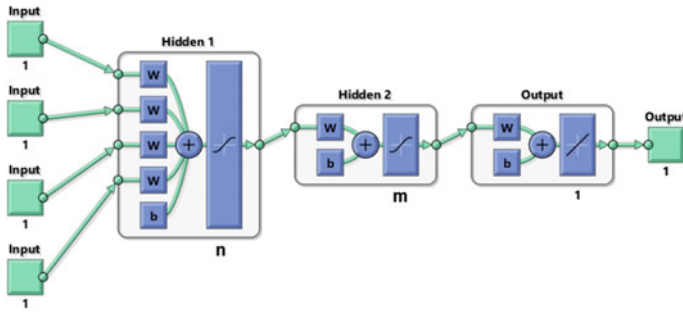


Fig. 7 General feed forward neural network scheme

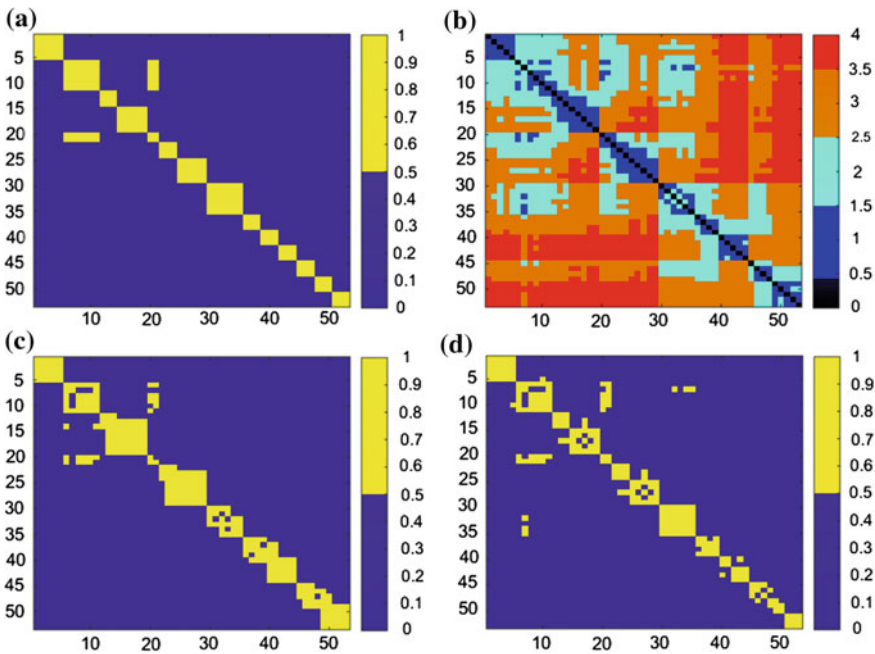


Fig. 8 Sibling identification: **a** reference, **b** GMM D_4-PC_1 , **c** DBScan-FC D_3-PC_1 , **d** FFNN

using the data described in Sect. 4.1. The best results (Fig. 8d) were achieved for the network with two hidden layers, where n was 159 and m was 2. Such a sizes might be justified by the number of marker dimensions ($53 \cdot 3$) and a small number of output classes (2).

3.5 Creating FBM

Grouping Siblings The machine learning described in Sect. 3.4 allow us to determine whether certain pair of markers can be siblings or not, although their raw results are error prone. To address the problem we propose the consensus grouping stage that creates final groups. The algorithm can be described as:

1. Selects the first, not assigned to any cluster marker and specify for him his siblings (candidates for group),
2. For each candidate we count the number of markers which regard it as their sibling,
3. We create a new group which includes all the markers that have been obtained the required minimum number of votes,
4. Markers which have not obtained the required number of votes returned to the set of unassigned,
5. Repeat steps 1–5 until all markers will be allocated to the group.

Parent Choice—we used the same as in the original algorithm. Each marker in the group proposes candidate markers (the closest as R^E out of group) to be a parent (Fig. 2b). Parent for the group is the group containing the most of markers that received the biggest number of votes.

4 Testing

4.1 Data

Algorithms have been tested on a set of 170 different calibration records acquired at 100 Hz using Vicon a MoCap system equipped with 10 T40 Vicon NIR cameras and Vicon Blade software in the human motion laboratory of the PJATK. The test sequences are a special kind of record where actor performs specific movements to cover its range of motion (ROM). Human body recordings were containing 53 markers located in standard marker configuration (Vicon 53 setup). In this scheme we placed two markers on each major joints (eg. the elbow, knee, ankle etc.), one marker between the joints, and a few markers placed in a critical anatomic areas such as pelvic, metatarsal, head, chest etc. Because the calibration recording can't be too long, the actors had a limited time for precise execution of all individual movements, although, the movement had also to be natural with no extreme velocities. The other subjects required custom marker setups. The data set consists of three main subsets:

1. Correct calibration of a healthy person (Fig. 9),
2. Calibration of a diseased person (after ischemic stroke or hip replacement) (Fig. 10),
3. Incorrect calibration of a healthy person (without some moves), (Fig. 11) and we included some extra recordings for comparison:

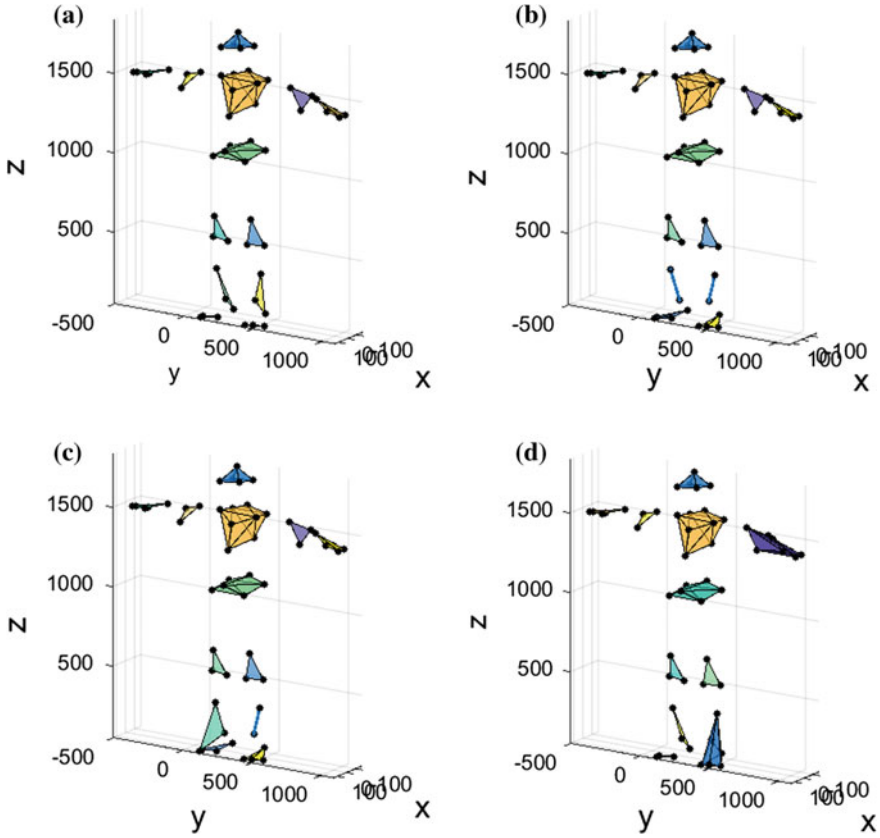


Fig. 9 FMB for HJ subject using: **a** genuine method, **b** DBScan-FC, **c** GMM Global, **d** FFNN

- (a) Record of the movements of the animal—a dog,
- (b) Record of the facial expressions—nonrigid subject,
- (c) Record of the movement of the object—three segment pendulum.

4.2 Evaluation Criteria

For the evaluation of the obtained FBM we assumed following criteria:

1. The numbers correctly/incorrectly assigned markers (Table 1a),
2. The time required to create FBM (Table 1b),
3. Subjective evaluation of experienced Motion Capture System workers.

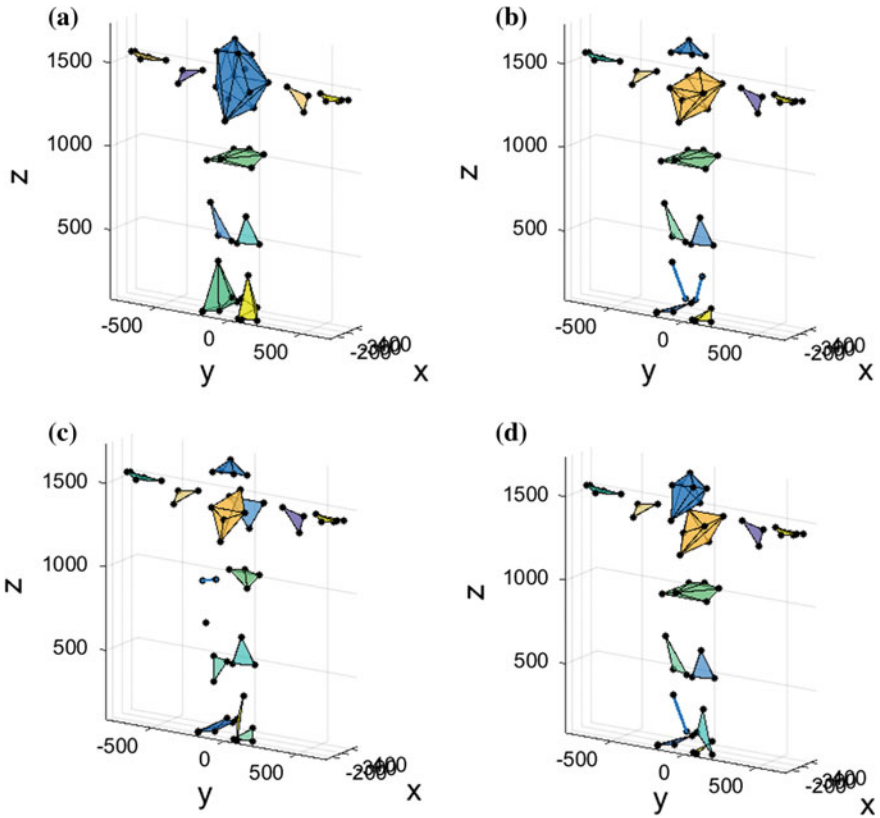


Fig. 10 FMB for JB subject using: **a** genuine method, **b** DBScan-FC, **c** GMM Global, **d** FFNN

4.3 Results

The quantitative results are demonstrated in Table 1. The time required to create a FBM (Table 1b) for each of the algorithms was comparable and less than half a second on a decent personal laptop (i7Core-4720HQ @ 2.6 GHz), so the time performance seems quite reasonably. The quality of the results in Table 1a demonstrates the best performance of the FFNN, but error for other approaches is comparable.

Looking for the exemplary cases—FBMs obtained with GMM were the worst ones—in most cases it tended to divide groups into small ones. The groups of markers located on pelvis (Figs. 10c and 11c) and chest (Fig. 10) were most frequently divided. The FBM obtained with the neural network in most cases were quite good. But in the case of a diseased person it incorrectly expanded group (Fig. 10d). FBMs based on the DBScan-FC also were quite good. This method had the worst results in the group of incorrectly performed calibrations—no movement in one or more joint (Fig. 11b).

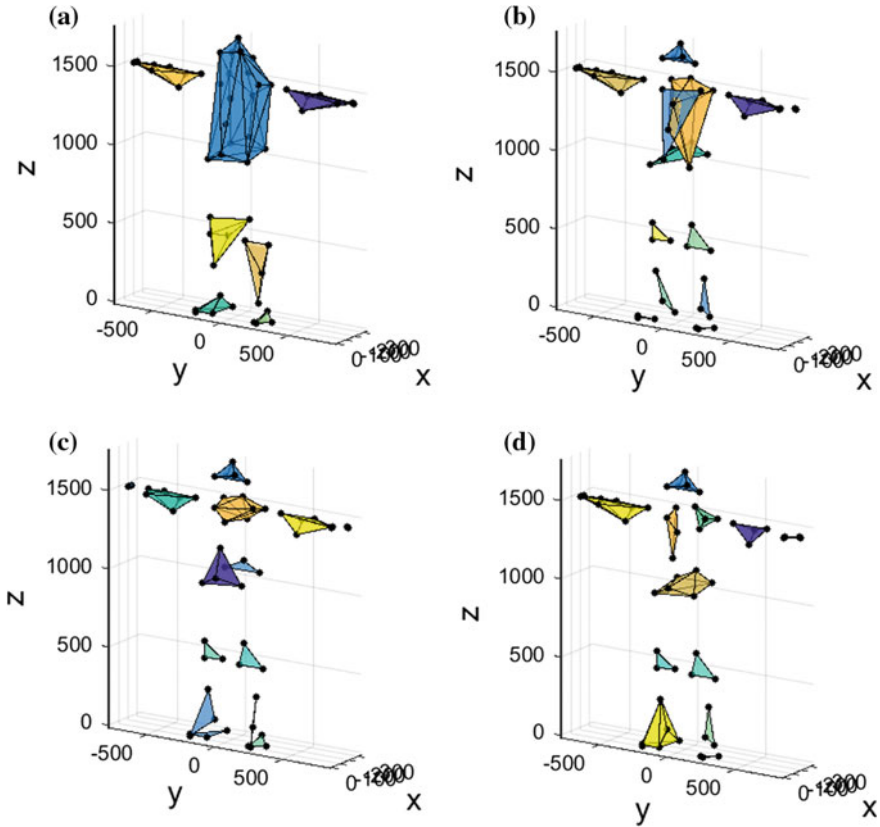


Fig. 11 FMB for MR subject using: **a** a genuine method, **b** DBScan-FC, **c** GMM Global, **d** FFNN

Although, in the group of a diseased person it identified different body segments properly. This algorithm was also the best one to identify of dog body parts.

Also for the non-human examples of a rigid body (not depicted here)—dog and pendulum the results were consistent with the results obtained for the human subjects, whereas the results for the human face sequence, which do not fulfill the rigid body requirement were rather meaningless as we expected.

We performed also brief analysis of susceptibility of the proposed approach to velocity. For the whole set of a complete ROMs for healthy subjects, we compared the length of average and peak gradient of a sequence with a number of erroneously classified markers. We analyzed Pearson linear correlation (PCC) and Spearman rank order correlation (SROCC) to discover nonlinear dependence. The obtained values (for mean gradient: PCC = -0.054 SROCC = -0.072 , for peak gradient: PCC = 0.12 SROCC = 0.2) suggest that velocity has no significant influence on the results.

Table 1 Number of erroneously grouped markers and time required to create FBM

(a) Error type	Method					
	GMM		DBScan-FC		FFNN	
	AVG	SD	AVG	SD	AVG	SD
Erroneously added to the cluster	11	4	10	5	6	6
Erroneously removed from the cluster	15	14	9	16	3	3
The total number of wrongly classified	26	15	18	10	9	8
(b) Time [s] required for						
Calculating the input matrix	0.4453	0.1543	0.3538	0.0049	0.0155	0.0005
Creating groups	0.0011	0.0010	0.0008	0.0002	0.000002	0.00004

5 Summary

Each of the proposed algorithms fared better or worse depending on the case. Each of the methods create reasonably looking hierarchical meshes. For healthy people correctly performing the calibration record best results were obtained for FFNN and DBScan-FC—these algorithms were comparable. The FFNN was better for improperly performing calibration, and DBScan-FC for diseased person. As each of the methods has some advantages we consider fusion of the results from each of them. The other promising directions for the body clustering we see to use kernel methods such as kLDA or kSVM.

Acknowledgments M. Pawlyta was supported by Virtual physiotherapist project (TANGO I WR) by The Polish National Centre of Research and Development. P. Skurowski was supported by the statutory research grant No. 02/020/BK_15/0061.

References

1. Belhumeur, P.N., Hespanha, J.A.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997)
2. Czekalski, P., Łyp, K.: Neural network structure optimization in pattern recognition. *Studia Inform.* **35**(4) (2014)
3. Daszykowski, M., Walczak, B., Massart, D.: Looking for natural patterns in data: part 1. density-based approach. *Chemometr. Intell. Lab. Syst.* **56**(2), 83–92 (2001)
4. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the em algorithm. *J. Roy. Statist. Soc. Ser. B* **39**(1), 1–38 (1977)
5. Figueiredo, M.A.T., Jain, A.K.: Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(3), 381–396 (2002)
6. Moeslund, T.B., Granum, E.: A survey of computer vision-based human motion capture. *Comput. Vis. Image Underst.* **81**(3), 231–268 (2001)

7. Moeslund, T.B., Hilton, A., Krüger, V.: A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.* **104**(2), 90–126 (2006)
8. Skurowski, P., Pawlyta, M.: Functional body mesh representation—a simplified kinematic model. In: *Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2014 (ICNAAM-2014)*, vol. 1648, p. 660008. AIP Publishing (2015)
9. Skurowski, P., Pawlyta, M.: Functional body mesh representation, a simplified kinematic model, its inference and applications. *Appl. Math. Inform. Sci.* **10**(1), 71–82 (2016)
10. Sokół, R., Polański, A.: Comparison of methods for initializing em algorithm for estimation of parameters of gaussian, multi-component, heteroscedastic mixture models. *Studia Inform.* **34**(1), 49–73 (2013)
11. Świtoński, A., Mucha, R., Danowski, D., Mucha, M., Polański, A., Cieślak, G., Wojciechowski, K., Sieroń, A.: Diagnosis of the motion pathologies based on a reduced kinematical data of a gait. *Electr. Rev.* **87** (2011)
12. Szczęsna, A.: Blobby quaternion model of human joint limits. In: *Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2014 (ICNAAM-2014)*, vol. 1648, p. 660008. AIP Publishing (2015)

Evaluation of Automatic Calibration Method for Motion Tracking Using Magnetic and Inertial Sensors

Ewa Lach

Abstract During the human motion tracking with miniature IMU sensors we cannot assume that the sensors are perfectly aligned with the segments to which they are attached. In order to find the relative rotation between the sensor and the segment, the sensor must be subjected to the step of calibration. The article concerns the evaluation of automatic calibration method searching for the sensor-to-segment orientation using linear acceleration of segments in the kinematic chain. Performance of the method is analyzed for simulated and real data.

Keywords Inertial measurement units · Sensor-to-segment calibration · Motion tracking

1 Introduction

Human motion tracking has applications in various domains, including virtual reality and biomedical engineering. Three-dimensional reconstruction of body movement is a promising tool for clinical evaluation and ambulatory monitoring. User-worn inertial sensors, also known as inertial measurement units (IMUs) has proven to be suitable for unrestrained tracking of body segments orientations due to the fact that they are small, light and completely self-contained. Generally used IMUs include accelerometers, gyroscopes and magnetometers. Due to their physical characteristics estimation results using this devices suffer from high measurement noise, incorrect scaling and biasing. Therefore, there are many studies discussing the development of efficient filters and algorithms exploiting the complementary properties of the sensors [14]. While the output signal from the IMU is based on the sensor coordinate system (CS), it is necessary to express the signal in the body segment. This requires a calibration method to determine the exact orientation of each of the sensors with respect to the segment.

E. Lach (✉)

Silesian University of Technology, Institute of Informatics, Gliwice, Poland
e-mail: Ewa.Lach@polsl.pl

The purpose of this work is to study an automatic calibration method proposed in [5] and validate its effectiveness for tracking human motion. The idea presented by authors in [5] was elaborated in this paper for the needs of the motion capture.

The remainder of the paper is organized as follows. First, in the Sect. 2 related works are presented. Then, in the Sect. 3 automatic calibration method is summarized. In the Sect. 4 experiments on simulated and real data and their results are presented and discussed. Finally, the Sect. 5 concludes the paper.

2 Related Work

Much research concerning IMU sensors has been conducted to reconstruct trajectories, or to estimate sensor orientations, typically by fusing a rate gyroscope, an accelerometer, a magnetometer and reference values such as Earth's gravity vector and magnetic field vector by different Kalman filters (for a comprehensive review on this topic, see [14]). There were also efforts to improve these orientation estimations by using the constraints arising from segments' attachment to each other [17].

Although, the human body lacks even surfaces and right angles, the human muscles and skin are non-rigid, many researchers in their work make an assumption that the sensors are perfectly aligned with the segments to which they are attached. Only a few calibration methods have been proposed for the IMU [1–3, 5, 7–9, 11–13, 15, 16, 18].

The most commonly applied methods, estimating orientations in which the sensors are mounted with respect to the body segments, are methods using predefined static calibration postures and predefined calibration motions.

In [1] it was assumed that when the human model is in the reference position (upright, facing north with arms along the body), the limb segment coordinate axes are aligned with the corresponding Earth-fixed axes. As a result, the orientation offset for each limb segment can be set to the inverse of the sensor orientation from the segment's reference position.

For calibration processes presented in [9, 13] segments don't need to align with Earth frame. It is sufficient to know segment's orientation in predefined position (T-pose or N-pose) to calculate orientation offset.

Because magnetometers are affected by ferrous compounds, many researchers have commented on their unsuitability for motion measurements, especially in home environments. To avoid errors caused by magnetic disturbances in [18] authors proposed calibration method using only the gravitational acceleration measured for two postures (standing upright and sitting flat with outstretched legs on the floor). The problem with that solution is necessity of setting limb segments with the axis of gravity vector.

Predefined motion calibration consists of performance of certain movements that are assumed to correspond to a certain axis. The measured orientation and angular velocity are used to find the sensor orientation with respect to the segment's functional axes [3, 11, 13].

In the reference calibrations the gravity data and angular velocity can be selected by the examiner (only a single data point is used to determine each axis) [11] or averaged for all the data measured during the rotation [3]. Another approach processes the angular velocity by the principle component analysis technique (PCA) to express the rotation axis in the IMU CS [8]. A different method applies singular value decomposition (SVD) to estimate axis of system of coordinates that are representative of real physiological motions. Moreover, associated with the estimates is a measure of the reliability of the computed axis, used to properly condition the regression algorithm Levenberg-Marquardt, which is applied to compute the transformation between each pair of systems of coordinates [12].

Often predefined static position is combined with predefined motions to increase efficiency of calibration [11]. In protocol ‘Outwalk’ presented in [2] in addition to the reference calibration appropriate placing of sensors on the segments is required. The protocol used in [13], on the other hand, solves a closed kinematic chain to refine results that have been obtained from a combination of calibration postures and predefined motion.

The mentioned methods make a major contribution to the quality of IMU-based measurements and are popular. However, both in calibration postures and calibration motions, the accuracy is limited by the precision with which the subject can perform the postures or motions.

Next type of calibration uses the knowledge of the attachments between the segments.

Knowing the orientation of the segments relative to each other and the global orientations of the sensors based on IMU data captured during the movement we can determine sensors orientations with respect to the segments [7]. Unfortunately, information about relative orientation between the segments isn’t always available.

The next works, also exploit the kinematic constraints, but they do not require any knowledge about the sensor’s positions or orientations. The information is extracted from the measurement data of almost arbitrary movements.

In [15, 16] least-squares methods for the estimation of joint axes and positions from gyroscopes and accelerometers measurement data are presented that exploit the kinematic constraints induced by hinge and spheroidal joints.

In [5] the non-linear optimization algorithm is applied to find relative orientation of the sensor from simulated and measured data for robot arm control. The authors proposed the use of linear acceleration of segments in kinematic chain to locate orientation offset.

The last two methods have the greatest potential. Human errors are not as critical for them as for the other methods. Additionally they could be implemented for continuous calibration during motion capture to prevent errors that result from sensors displacement during the movement.

3 Automatic Calibration Method

In this paper the automatic calibration method proposed in [5], based on linear acceleration data analysis, is evaluated for its effectiveness for human motion tracking. The calibration method involves creating a model of a moving object with simulated sensors on each model's segment. The algorithm compares simulated data (linear acceleration) with measured data and by the means of an optimization algorithm determines the offset between the sensor CS and anatomical CS. Simulated accelerometer generates signal as a second derivative of position in simulated sensor CS.

The evaluated algorithm consists of the following steps:

1. Recording object's motion data.
2. Signal preprocessing.
3. Setting sensor orientation relative to the segment.
4. Moving the model with the measured orientation adjusted by the sensor offset. Obtaining the simulated data from the model.
5. Calculating the distance (the error) between the measured data and the data derived from the model.
6. Returning the offset if the resulting error is small enough or the last iteration of the algorithm was completed, else continuing with step 7.
7. Changing the offset to minimize the error by the means of the optimization algorithm. Continuing with step 3.

In the implemented algorithm signal preprocessing include:

- the estimation of the sensor orientation in global CS,
- removal from the accelerometer signal the acceleration caused by gravity,
- noise reduction of measured linear acceleration by a low pass filter (81 point moving average filter (MAF)).

The distance function is:

$$d = \sqrt{((a_s - a_M) \cdot mp)^2}, \quad (1)$$

where d is an error for a single sample, a_s is linear acceleration due to motion, a_M —acceleration from simulated sensor, mp —three-element vector parameter comparing monotonicity of simulated signal and measured data in the corresponding intervals. Values of parameter mp were determined experimentally. They are equal to 1, when monotonicity of compared data is the same, 2 if they are opposed, 1.25 if measured signal is constant, 1.6 if simulated signal is constant.

As we are looking for the smallest possible error the calibration can be described as optimization problem (where vector of distances of all the samples is minimized), which is solved by using the Levenberg-Marquardt algorithm (LM). The searched orientations are Euler angles measured in radians. For three segments we are looking for nine values. At the first iteration of the algorithm it is assumed that the sensor is

perfectly aligned with the segment. The value of the differentiation step (ds) of The Levenberg-Marquardt algorithm has been determined experimentally at 0.001. The behavior of the LM algorithm has been also studied for its higher ($ds = 0.01$ in test T_LM01) and lower ($ds = 0.0001$ in test T_LM0001) value. The results can be seen in Table 2. In the end the value 0.001 has been selected.

4 Experiments

The evaluation of performance of presented calibration method was done on the basis of the averaged deviations between true and estimated offset orientations for simulated data and between synchronized orientations of segments captured by Vicon system and IMU sensors for real sensors data. Statistical analyses comprised RMS, mean with standard deviation (SD) as well as index DI corresponding to the geodesic distance between two quaternions: estimated by calibration q_C and the true orientation q , on the hypersphere S^3 : $DI = 2 \cdot \arccos(|q_C \cdot q|)$ [4]. We use those three measures because the first two are the most popular in publications about motion capture and the last measure should better represent the difference due to the fact that a quaternion describes rotation on multiple axes at once.

4.1 Calibration of the Simulated Data

At first, we tested the proposed calibration on synthetic data for which we knew the orientation offset between the segment and the sensor attached to it. The three-dimensional virtual model of three-segment pendulum has been constructed (Fig. 1b). Segments are connected by movable joints with three rotational degrees of freedom (DOF). Simulated sensors were mounted at the segments. In animation framework fACT [10], which was used for simulation, Euler angles are measured in radians. However, to improve the understanding, in this paper all angles are presented in degrees. Each test was repeated 50 times. Standard test parameters values are presented in Table 1. They are the result of the experiments presented below. They were applied in test T_std. First the calibration method was tested for different number of sensors, which means different number of searched values. We applied evaluated algorithm to three-segments (test T_S3), two-segments (test T_std) and one-segment (test T_S1) model. In subsequent experiments, we searched by means of LM algorithm for, respectively, 9, 6 and 3 values.

In the Levenberg-Marquardt algorithm the user has to provide an initial guess for the parameter vector. In cases with multiple minima, as is in that case, the algorithm converges to the global minimum only if the initial guess is already somewhat close to the final solution. For the initial iteration vector of zeros seems the most natural (the sensor is perfectly aligned with the segment) for orientation angles.

Fig. 1 The pendulum (a) and its model (b)

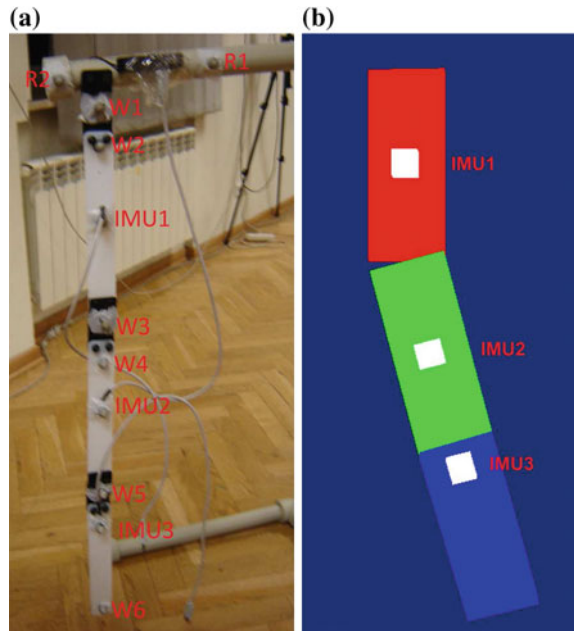


Table 1 Standard test parameters

Parameter	Value
Number of moved segments	2
Number of samples	1000
Offset range	$\langle -25^\circ, 25^\circ \rangle$
Search offset range	$\langle -25^\circ, 25^\circ \rangle$
Noise	Absent
Acceleration due to gravity	Absent
Acceleration amplitudes	16 m/s^2
LM differentiation step	0.001

For the three tests the impact of the value of the offset on the quality of the calibration was checked. In the test T_std for each axis, the rotation offset is within the range $\langle 25^\circ, 25^\circ \rangle$. For test T_O45 the interval $\langle -45^\circ, 45^\circ \rangle$ have been used and for test T_O90— $\langle -90^\circ, 90^\circ \rangle$. The limits on values searched by LM algorithm were, also, added corresponding to offset intervals.

In addition, we tested how the calibration method will behave for a larger search domain of LM algorithm. In test T_L90 the found values cannot extend beyond range $\langle -90^\circ, 90^\circ \rangle$, in test T_L180—beyond range $\langle -180^\circ, 180^\circ \rangle$.

The generated object motion is characterized by high values of acceleration amplitudes of about 16 m/s^2 . In test T_A9 the acceleration amplitudes were reduced to 9 m/s^2 , in test T_A5 to about 5 m/s^2 and then in test T_A2 to about 2 m/s^2 . The tests aim was founding out how the change of the nature of the movement will affect obtained results.

Because accelerometers are characterized by high measurement noise, another tests checked how calibration algorithm will behave with noise added to simulated data. To accelerometer signal were added a Gaussian distributed random values with mean 0 and standard deviation 0.05 for test T_N005 and 0.1 for test T_N01. Since signal preprocessing include noise reduction additional test T_N005NF was performed without low pass filter applied to linear acceleration.

In the test T_G the gravitational acceleration was added to generated signal with vector of gravity $g^N = [0, 0, -9.81]^T$. The influence of this measure on simulated calibration was verified. However, the results obtained for simulated data may not correspond well with the values acquired for measured data due to the fact that in measured accelerometer signal linear acceleration is combined with gravitational. The removal of gravity from the measured accelerometer signal generates an additional error as a result of the inaccuracy of the removal process. In simulated calibration errors caused by the removal of gravity from the accelerometer signal are not taken into account. Results obtained for experiments on simulated data, described above are presented in Table 2. In the test T_S1 a solution for one segment has been searched. We can see in Table 2 that despite the fact that the search algorithm had the least values to find the error is one of the largest of all obtained in all the tests. It is believed that the reason for that is the large number of possible IMU offsets generating similar signal of linear acceleration. Adding second segment (test T_std)—creating kinematic chain—greatly improved the results for the first sensor. The second segment has imposed restrictions on the first sensor orientation. It must also be noted that the error for the second sensor in test T_std is again large, because the effect of kinematic constraints does not exert the same force on the last segment in the chain as on the first and middle one. However, there is some influence, because deviation index for second sensor is smaller than the DI for test T_S1. Adding the third segment (test T_S3) imposes constraints on the first and the second sensor. But more data is associated with poorer matching for the first segment (DI=0.72 for the first sensor in the test T_std increased to 1.5 for the T_S3 test).

Based on the above results the use of the hierarchical calibration is suggested. For a single sensor we will be looking for an offset based on the signal from the two sensors. Established offset for the first sensor can be used as a constant in the search for the offset of the second sensor.

Another factor having a strong influence on the quality of the calibration is the amount of the orientation offset. The greater the offset, the less chance of finding the correct solution. The error for test T_O90 is so large that the result suggests that we cannot use this calibration method for big offsets at all. In this situation the calibration should be used only for small offsets. It shows limitation of the evaluated method but it is not a critical issue. If we mount sensor with some care the offsets shouldn't be very large.

Table 2 The averaged deviations (in degrees) between true and estimated offset orientations for calibration of simulated data

Test (segment)	Mean (SD)			RMS	DI
	Roll	Pitch	Yaw		
T_S1 (1)	2.47 (1.68)	0.36 (0.24)	19.03 (12.88)	19.20	19.09
T_std (1)	0.02 (0.05)	0.00 (0.01)	0.72 (1.79)	0.72	0.72
T_std (2)	2.02 (1.74)	0.31 (0.23)	15.55 (13.35)	15.69	15.31
T_S3 (1)	0.02 (0.04)	0.02 (0.08)	1.52 (2.93)	1.52	1.50
T_S3 (2)	0.04 (0.15)	0.02 (0.04)	0.53 (1.14)	0.54	0.55
T_S3 (3)	13.57 (10.48)	2.59 (2.70)	7.37 (6.11)	16.01	16.59
T_O45 (1)	0.05 (0.28)	0.01 (0.03)	2.09 (9.79)	2.09	2.14
T_O90 (1)	1.04 (2.53)	11.01 (37.40)	36.17 (57.44)	40.91	39.71
T_L90 (1)	0.06 (0.23)	0.05 (0.23)	2.42 (9.96)	2.42	1.29
T_L180 (1)	0.03 (0.11)	0.01 (0.02)	1.11 (3.60)	1.11	1.13
T_A9 (1)	0.03 (0.13)	0.01 (0.02)	1.18 (5.28)	1.18	1.18
T_A5 (1)	0.04 (0.10)	0.01 (0.02)	1.23 (2.74)	1.23	1.25
T_A2 (1)	0.07 (0.17)	0.02 (0.06)	1.48 (3.34)	1.48	1.50
T_N005NF (1)	0.10 (0.31)	0.01 (0.03)	3.13 (7.72)	3.14	3.16
T_N005 (1)	0.01 (0.03)	0.01 (0.01)	0.49 (1.16)	0.49	0.50
T_N01 (1)	0.02 (0.03)	0.01 (0.01)	1.00 (1.76)	1.00	0.97
T_LM0001 (1)	0.02 (0.09)	0.00 (0.03)	0.69 (2.99)	0.69	0.70
T_LM01 (1)	0.11 (0.16)	0.03 (0.04)	4.20 (5.63)	4.20	4.29
T_G (1)	0.04 (0.12)	0.01 (0.04)	1.28 (4.01)	1.28	1.30
T_G_O90 (1)	2.22 (7.87)	15.82 (47.67)	14.66 (37.81)	23.39	20.54

In the tests T_L90 and T_L180 a broadening of search domain has been studied for small offsets (in interval range $\langle -25^\circ, 25^\circ \rangle$). The obtained errors imply that on average we get better results with limits. It should be added that the broadening of the search domain results in the appearance in the 50 samples of one or several solutions with the error level of several tens to 180° . For example, after removal of two solutions from 50 for test T_L90 the error DI was equal to 0.42° .

Added noise did not increase errors (tests T_N005 and T_N01). This can be attributed to the noise reduction filter used in tests. When the MAF was removed in T_N005NF the error DI raised to 3.16° .

The high values of acceleration amplitudes means more information for search algorithm, so if it is possible, to provide better results, the calibration motion should include fast motion (see T_std, T_A9, T_A5, T_A2 in Table 2).

The inclusion of gravity to the generated signal resulted in small deterioration of the calibration (test T_G). Additional data obscured the signal. In the course of the experiments with real data it should be checked which generates greater errors: gravity in accelerometer signal or its removal.

Additionally, in all conducted experiments it was observed that the results for yaw rotation were always worse than for the pitch and roll. The error for yaw offset was on average 50 times higher. This observation demands more study. It is possible that we can use it to improve calibration algorithm.

4.2 Calibration of the Real Sensor Data

To test effectiveness of evaluated calibration method on real sensor data the pendulum build with three segments connected by movable joints and its virtual model (Fig. 1) were used. To each segment an IMU sensor built at the Silesian University of Technology, Department of Automatic Control and Robotics [6] was attached. These IMU sensors were marked as IMU1, IMU2, IMU3 (Fig. 1a). As a reference, data from the optical system of motion capture (Vicon system) was used. On the pendulum markers, marked as R1, R2, W1, W2, W3, W4, W5 and W6 were attached. The IMU sensors and the Vicon system worked with the same frequency of 100Hz. The recording had a length of 10000 samples, from which we remove 7000 samples without linear acceleration (the pendulum wasn't moving). The recorded movement had high values of acceleration amplitudes (about 20m/s²), but also the motion was mostly in one axis (Fig. 2). The estimation of the sensor orientation in global CS was accomplished by complementary filter delivered with the sensor.

Tests on simulated data have shown how sensitive to a variety of factors is evaluated calibration algorithm. Tests on real data only confirmed it.

4 experiments were performed concerning calibration with real sensor data. Tests T_{rgo} and T_{ro} have large orientation offsets, when T_{rg}, and T_r have small offsets. The acceleration due to gravitation was removed from acceleration signal in tests T_{ro} and T_r and it was left in tests T_{rgo} and T_{rg}.

We need to compare orientation data streams from Vicon Nexus system and IMU sensors, with different reference frames and measured with separate timers (with the same frequency). As a result, the data must be synchronized (in the time domain) and transformed to the same reference frame. As a reference frame body frame was used. To synchronize the data we find the time offset between the two signals for which the distance between two signals is minimal. Time synchronization is performed on angular velocity of the body. For the Vicon system angular velocity must be calculated. For this purpose we use equation:

$$\omega_{Vicon} = 2 \cdot q_{Vicon}^{-1} \otimes \frac{dq_{Vicon}}{dt} \tag{2}$$

where q^{-1} is the inverse of q and \otimes denotes quaternion multiplication.

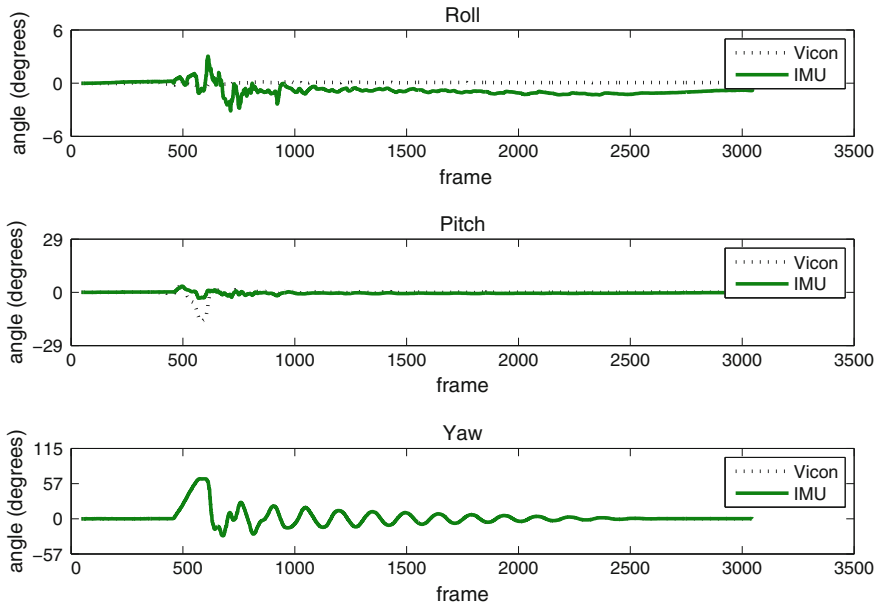


Fig. 2 Estimated Euler angles of segment 1 from calibrated IMU sensor and Vicon system for test T_r

Table 3 The averaged deviations (in degrees) between reference data from Vicon system and sensor IMU1 (after calibration)

Test (segment)	Mean (SD)			RMS	DI
	Roll	Pitch	Yaw		
T_r (1)	0.86 (0.43)	0.85 (1.90)	11.06 (15.99)	11.37	9.07
T_rg (1)	1.50 (3.17)	2.58 (6.53)	10.50 (15.12)	11.51	9.50
T_ro (1)	20.72 (21.53)	6.78 (7.28)	26.22 (31.29)	34.71	33.03
T_rgo (1)	2.30 (2.57)	4.68 (8.14)	8.28 (13.82)	10.82	10.55
T_mc (1)	0.86 (0.45)	0.86 (1.90)	11.07 (16.00)	11.38	9.06

Results of experiments (Table 3) confirm that the LM algorithm has trouble finding the right orientation for big offsets. When the sensor was turned about 180° in respect to segment CS for one of the axis obtained errors were very significant (T_ro). Interestingly, leaving gravity in accelerometer signal improved the results (T_rgo). The big difference in the observed errors allows to speculate that the presence of gravity may limit the number of minimums. For small differences in orientation of the sensor relative to the orientation of the segment positive aspect of the gravity decreases. Therefore, we observe a slight deterioration in results in the test with gravity (T_rg) in relation to the test without gravity (T_r).

To verify an idea that gravitational acceleration improves results for big offsets additional test T_G_O90 on simulated data was performed. The offsets were generated in the interval $(-90^\circ, 90^\circ)$. The DI error went down about 19° , which confirms observed results in real data. In Fig. 2 the results from test T_r (an estimation of orientation of segment 1) are presented in the form of three Euler angles. As a reference the angles from Vicon system are also shown. The data has been synchronized and converted to the same frame.

Because the error DI for test T_r (equal to about 9°) seemed still rather big the results from presented calibration method with the results obtained from optical system (T_mc) were compared. The error is only slightly smaller, which leads to believe that the error caused by calibration in T_r is at a level of 0.01° , rather than 9° . This 9° is caused by inaccuracy of sensors and used complementary filter.

It is important to remember that the errors calculated for real data are the averaged distances between synchronized orientations of segments captured by Vicon system and IMU sensors. They are not the distances between true and estimated offsets like for simulated data. To show the difference between these values the orientation offsets from test T_mc and T_r were compared. The errors $RMS = 3.15$ and $DI = 3.22$ were obtained. As we can see in this case the error in offset of about 3° corresponds to error of 0.01° for motion data.

5 Summary

In this paper the evaluation of automatic calibration method is presented. The results are shown for simulated and real data. In 24 experiments many parameters (e.g.: number of segments, value of offset, noise) influencing quality of the calibration process were tested. The results shows that the presented calibration method, with some limitation can be used to find sensor-to-segment orientation. The next step in research will be to perform more experiments with real data. We need to test the method for motion in all three axis with a human skeleton model.

Acknowledgments This work is part of the General Statutory Research Project 02/020/BK_15/0061 conducted at the Institute of Informatics, the Silesian University of Technology.

References

1. Bachmann E., McGhee R., Yun X., Zyda M.: Inertial and magnetic posture tracking for inserting humans into networked virtual environments. In: ACM Symposium on Virtual Reality Software and Technology (2001)
2. Cutti, A., Ferrari, A., Garofalo, P., Raggi, M., Cappello, A., Ferrari, A.: 'Outwalk': a protocol for clinical gait analysis based on inertial and magnetic sensors. *Med. Biol. Eng. Comput.* **48**, 17–25 (2010)

3. Favre, J., Aissaoui, R., Jolles, B., deGuise, J., Aminian, K.: Functional calibration procedure for 3D knee joint angle description using inertial sensors. *J. Biomech.* **42**, 2330–2335 (2009)
4. Gramkow, C.: On averaging rotations. *J. Math. Imaging Vis.* **15**(1–2), 7–16 (2001)
5. Hoffmann, J., Brüggemann, B., Krüger, B.: Automatic calibration of a motion capture system based on inertial sensors for tele-manipulation. In: proceedings of 7th International Conference on Informatics in Control, Automation and Robotics (ICINCO), Funchal, Madeira, Portugal (2010)
6. Jedrasiak, K., Daniec, K., Nawrat, A.: The low cost micro inertial measurement unit. In: *Industrial Electronics and Applications (ICIEA)*, pp. 403–408. IEEE (2013)
7. Kang, D., Jung, Y., Park, A., Kim, J.: Human body motion capture system using magnetic and inertial sensor modules. In: *International Universal Communication Symposium* (2011)
8. Kim, J., Ryu, M., Choi, H., Yang, Y.: Anatomy calibration of inertial measurement unit using a principle component analysis. *Int. J. Bio-Sci. Bio-Technol.* **5**(6), 181–190 (2013)
9. Kok, M., Hol, J., Schon, T.: An optimization-based approach to human body motion capture using inertial sensors. In: 19th World Congress The International Federation of Automatic Control Cape Town, South Africa (2014)
10. Lach, E.: FACT—animation framework for generation of virtual characters behaviours. In: *International Conference on Information Technology*, pp. 325–328. IEEE (2008)
11. Luinge, H., Veltink, P., Baten, C.: Ambulatory measurement of arm orientation. *J. Biomech.* **40**(1), 78–85 (2007)
12. Ricci, L., Formica, D., Sparaci, L., Lasorsa, F., Taffoni, F., Tamilia, E., Guglielmelli, E.: A new calibration methodology for thorax and upper limbs motion capture in children using magneto and inertial sensors. *Sensors* **14**, 1057–1072 (2014)
13. Roetenberg, D., Luinge, H., Slycke, P.: Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors. *XSENS TECHNOLOGIES—version 3.4.2013* (2013)
14. Sabatini, A.: Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors* **11**(2), 1489–1525 (2011)
15. Seel, T., Raisch, J., Schauer, T.: IMU-based joint angle measurement for gait analysis. *Sensors* **14**, 6891–6909 (2014)
16. Seel, T., Schauer, T., Raisch, J.: Joint axis and position estimation from inertial measurement data by exploiting kinematic constraints. In: *Proceedings of the IEEE International Conference on Control Applications, Dubrovnik, Croatia*, pp. 45–49 (2012)
17. Szczęsna, A., Prusowski, P., Słupik, J., Pęszor, D., Polański, A.: Evaluation of improvement in orientation estimation through the use of the linear acceleration estimation in the body model. *Man–Machine Interactions*, vol. 4, pp. 377–387. Springer International Publishing (2016)
18. Takeda, R., Tadano, S., Natorigawa, A., Todoh, M., Yoshinari, S.: Gait posture estimation using wearable acceleration and gyro sensors. *J. Biomech.* **42**, 2486–2494 (2009)

The Influence of Uncertainty in Body Segment Mass on Calculated Joint Moments and Muscle Forces

Magdalena Żuk and Celina Pezowicz

Abstract Musculoskeletal models and computer simulations have been extensively developed in recent years. The purpose of present study was to investigate the effect of uncertainty in segment masses on joint torques and muscle forces. Segment mass uncertainties estimated on the basis of Dempster measurement have been introduced to the musculoskeletal model Gait 2392. Inverse dynamics, static optimization and CMC algorithm have been performed using OpenSim software. The segment mass uncertainty has limited effect on the net joint moments. However, the error propagation in proximal direction have been observed and the greatest impact was observed for residual moments and forces acting on pelvis. The greatest effect on muscle forces can be observed for both static optimization and CMC results. Furthermore, considerable differences can be seen between static optimization and CMC results. Therefore, an accurate estimation of muscles forces and load condition in musculoskeletal system is still posing a challenge.

Keywords Musculoskeletal modeling · Joint moment · Muscle force · Sensitivity analysis · Body segment parameters

1 Introduction

Knowledge of internal loads in human musculoskeletal system can be of great importance for the clinical assessment in neurology and orthopaedic, design of prostheses or injury prevention. Direct measurements of muscle forces are invasive and generally not feasible in clinical practise. Therefore computer simulation techniques have been extensively developed. Many musculoskeletal models from different anatomical data sets have been proposed in recent years [5].

M. Żuk (✉) · C. Pezowicz
Wrocław University of Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: magdalena.zuk@pwr.wroc.pl
URL: <http://www.biomech.pwr.wroc.pl>

Mostly, musculoskeletal system is modelled as the multibody system composed of rigid bodies connected by mechanical joints. This multibody systems is actuated by set of actuators which represent muscles. The inverse dynamic technique is used for determination of the net joint torques. Due to the redundancy of the muscle system, indication of the individual muscle force requires solution of the force distribution problem using optimization methods. However validation of dynamics simulation results is still limited.

Nowadays, there is a trend to increasing complexity of the model together with the introduction of subject-specific data [2, 6]. The so called personalised musculoskeletal modelling have emerged together with the availability of new measuring techniques. Anatomical data is incorporated in dynamic models [19, 22], as well as kinematics model [21, 26, 27].

Body segment parameters (BSP), among many others model parameters, can be personalised using such method as medical imaging, including dual-energy X-ray absorptiometry, CT, MRI, gamma-scanning, surface scanning or estimated using different techniques. The justification of such measures in musculoskeletal modelling can be evaluated using the sensitivity analysis.

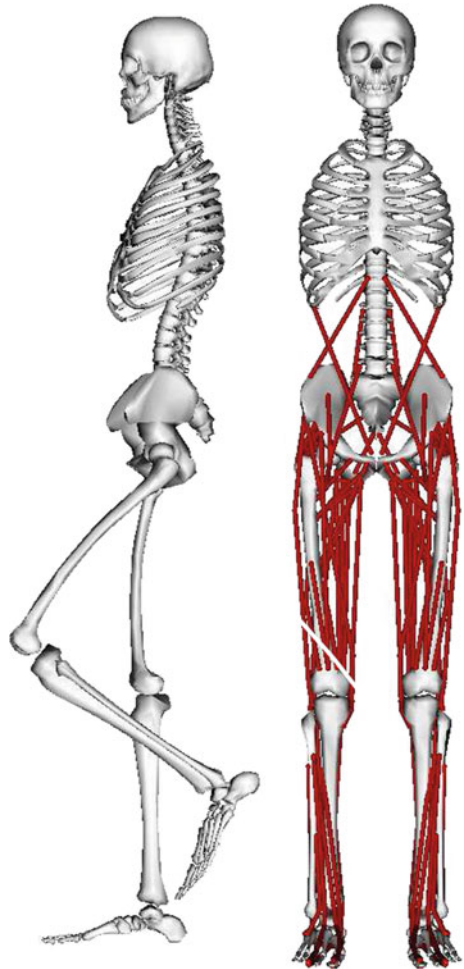
There have been many studies concerning the effect of selected body mass parameters on the net joint moments [7, 9, 10, 14–17, 20] and definitely less on individual muscle forces [24, 25]. Some authors content that the the effect of variability in body segment parameters on calculated joint moments is high [10, 16, 18], while others report only small impact [7, 13, 15, 17, 20, 25]. Nevertheless, perturbation of body mass parameters could better influence on force of selected muscles [24, 25]. Sensitivity analysis has been performed for very different models (often proposed by authors), including plantar [9, 14, 15, 18, 24] and three-dimensional models [16, 17, 20, 25], whole body models [14, 17, 18] and partial models. Furthermore, authors perturbed different body mass parameters (segment mass, centre of mass location, inertia tensor) individually or simultaneously. Range of perturbation usually seems to be arbitrary, although in some papers different method of BSP identification have been taken into account and compared [9, 16]. There are, according to the authors knowledge, no study, which reliably simulate possible errors in segment mass estimation and investigate the effect of these errors on calculated joint torques as well as on muscle forces.

The purpose of the present study was to investigate the effect of uncertainty in segment masses on joint torques and muscle forces with the following assumptions: the mass uncertainty is estimated on the basis of direct measurements, segment masses are perturbed together with resulting inertia tensors, the total body mass is constant, three-dimensional open access musculoskeletal model together with available simulation techniques are tested.

2 Methods

The sensitivity analysis was performed for open access computer model Gait2392 [3], 23 degrees of freedom model consists of 12 rigid segments: pelvis, torso, left and right thigh, shank, talus, calcaneus and toes and 76 muscles represented by 92 actuators (Fig. 1). Experimental data (trajectories of 49 markers and force data) collected by John et al. [8] on a subject walking (man, 72.6 kg) on the treadmill (1.13 m/s) was used for the simulation. Dynamic simulations were performed using OpenSim 3.0.1 software. The general model was scaled using measurements of marker positions in static pose.

Fig. 1 Model Gait2392 [3]



Segment mass estimation method implemented in OpenSim is a proportional scaling base on additionally scaled body mass parameters from Anderson et al. [1]. Those, in turn, were obtained for five living subjects using indirect method proposed by McConville et al. [12]. Segment mass uncertainty was assessed using data reported by Dempster [4]. Dempster introduced detailed anatomical data (including segment masses, mass centres, inertia tensors) collected from eight complete cadavers. For each subject, masses of segments were estimated using method implemented in OpenSim. Then, estimated and directly measured masses were compared, and average relative errors were calculated. Those uncertainties have been introduced to the initial model.

The manner of perturbation process is presented in Fig. 2. The modification has involved change of segment mass and owing inertia tensor (resulting from change in mass). Four modified models have been prepared, one with modification within three segments simultaneously (foot, shank, thigh) and three models with the modification within each segment independently. Together with those perturbations, torso mass and inertia tensor have been also modified in order to preserve constant whole body mass.

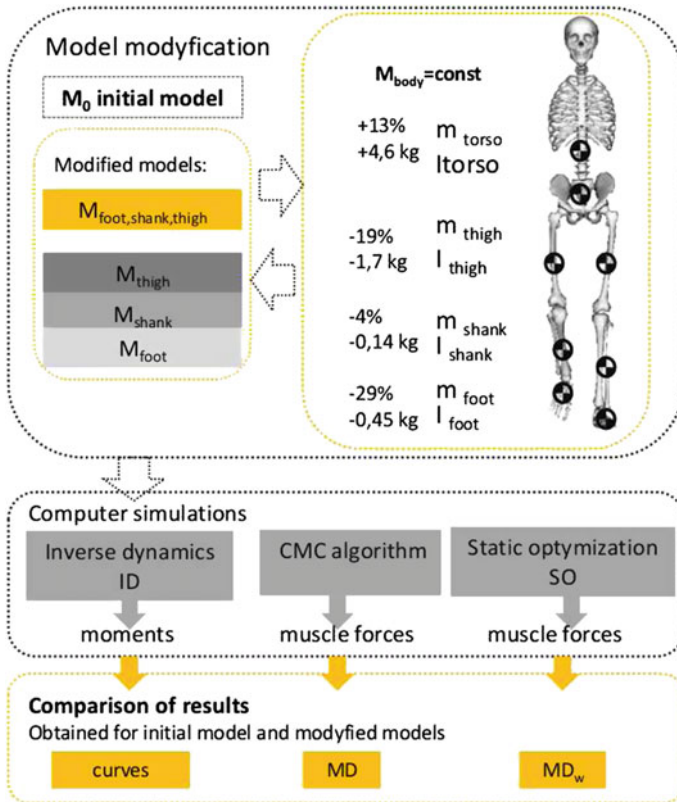


Fig. 2 Diagram showing model perturbation process and data analysis

The net joint torques were calculated for each of 5 models using inverse dynamic method described by Kuo [11]. Muscle forces were calculated using two different methods implemented in OpenSim: static optimization and CMC algorithm (*Computed Muscle Control*) described in [23]. Joint torques and muscle forces corresponding with one gait cycle were interpolated and normalized to 100 point using Matlab. Results of dynamic simulations obtained for initial and perturbed models have been compared.

In order to quantitative comparison, following parameters have been calculated:

$$MAV = \frac{1}{n} \sum_{i=1}^n \|M_{0_i} - M_{p_i}\|, \quad (1)$$

$$MAV_w = \frac{MAV}{\|M_0\|}, \quad (2)$$

$$MD = \frac{1}{n} \sum_{i=1}^n (M_{0_i} - M_{p_i}), \quad (3)$$

$$MD_w = \frac{MD}{\|M_p\|}, \quad (4)$$

where M_{0_i} is the value corresponding to the initial model, M_{p_i} is the value corresponding to the perturbed model at time i , $\|M_0\|$ is an absolute value of M_0 , averaged over the gait cycle.

3 Results

Estimated uncertainty introduced to the initial model were: reduction of the thigh weight of 19 % (1.7 kg), reduction of the shank weight of 4 % (0.14 kg), reduction the foot mass of 29 % (0.45 kg). The magnitude of these error is typical for segment mass identification using different methods [15, 16].

The inverse dynamics solutions are five sets of the net joint moments and residual forces and moments. Generally, changes introduced to the model have the limited impact on joint moments. The courses of torques in gait cycle have been preserved. The highest differences were in joint moments in sagittal plane, comparison of MAV_w is presented in Fig. 3. Values of MAV do not exceed 3 Nm. Larger discrepancies are for residual forces and moments. Residual moments in sagittal plane vary from 19 % for the thigh modification to 35 % for the foot modification. In the case of residual forces, discrepancies are even higher, there are $MAV_w = 80\%$ ($MAV = 10\text{ N}$) for the force in the sagittal plane.

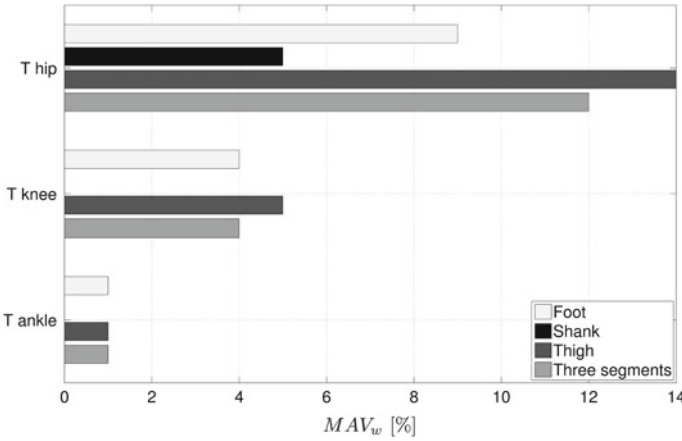


Fig. 3 Relative differences (MAV_w parameters) in sagittal joint moments (T_{ankle} -ankle dorsi/plantar flexion, T_{knee} knee flexion/extension, T_{hip} hip flexion/extension) calculated for initial (M_0) and four modified models (M_{foot} , M_{shank} , M_{thigh} , $M_{foot,shank,thigh}$) averaged over the gait cycle

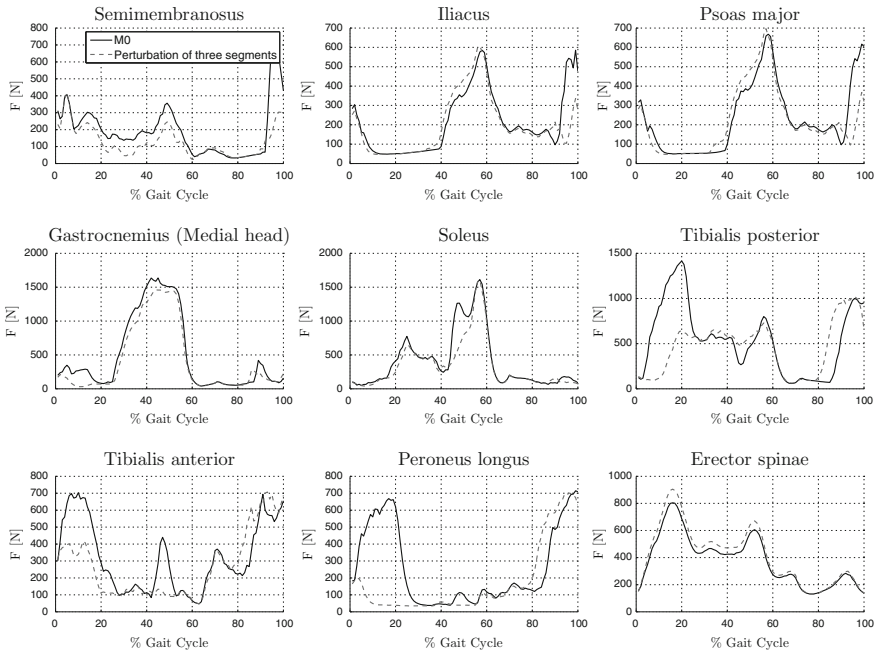


Fig. 4 Forces of selected muscles estimated using CMC algorithm during gait cycle, calculated for the initial model (M_0) and the perturbed model ($M_{foot,shank,thigh}$)

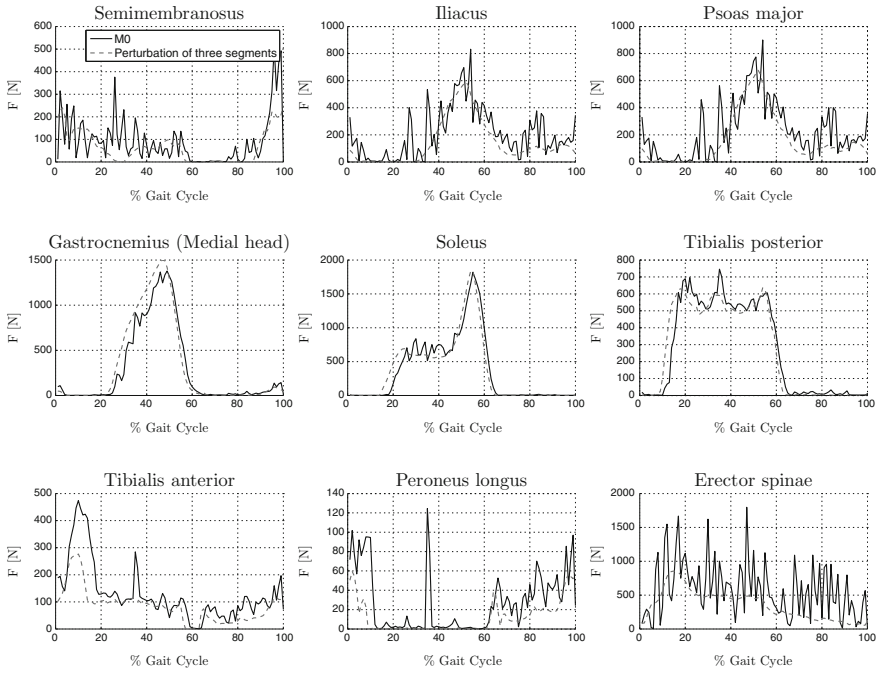


Fig. 5 Forces of selected muscles estimated using static optimization technique (SO) during gait cycle, calculated for the initial model (M_0) and the perturbed model ($M_{foot,shank,thigh}$)

The influence of model perturbations on individual muscle is presented for selected muscles for which this influence is greatest. The courses of muscle forces obtained for initial M_0 and perturbed model $M_{foot,shank,thigh}$ (modification in three segments) are presented in Fig. 4 for CMC and in Fig. 5 for the static optimization.

In the case of muscle forces obtained using CMC algorithm, for some muscles force values are changed, but course is preserved (medial head of gastrocnemius, psoas major, iliacus). For others, the muscle force is significantly reduced in the certain time span (tibialis posterior, tibialis anterior, peroneus longus), differences are even higher than 500 N. Static optimization results for initial model are irregular curves with large fluctuation for selected muscles. While for perturbed model, force curves are smoothed for those muscles. Values of MD parameters for each model are presented in bar plots (Fig. 6).

4 Discussion

Obtained joint moments are consistent with those from literature [9, 16, 20]. Segment mass and inertia tensor slightly affect the joint moments, which is confirmed by other authors [7, 13, 15, 17, 20, 25]. The certain trend can be seen, the more

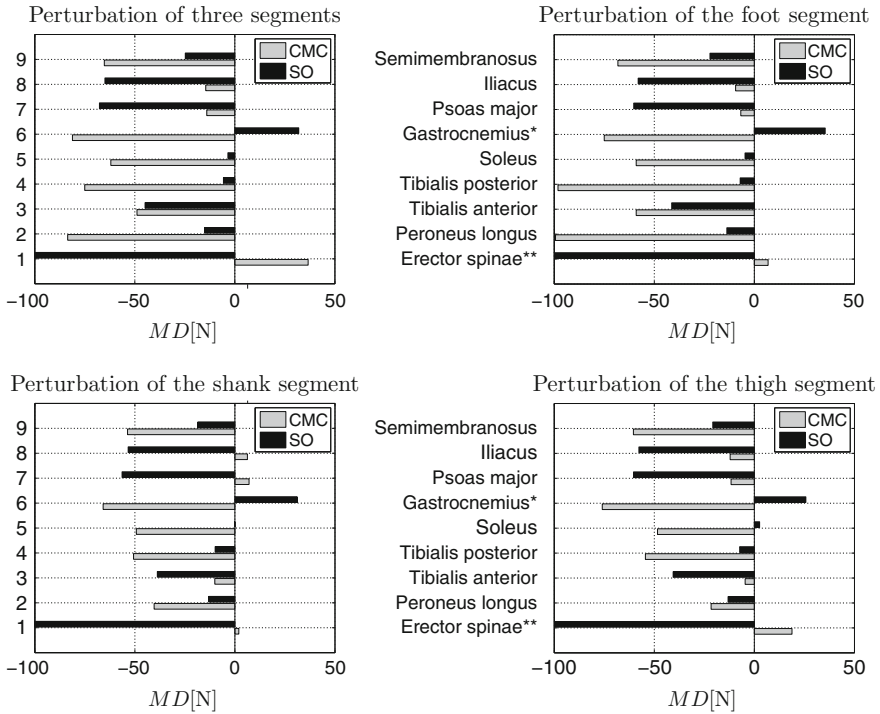


Fig. 6 Differences in muscles forces (MD parameters) calculated using the initial (M_0) and four modified models ($M_{foot,shank,thigh}$) averaged over the gait cycle. *Medial Head of Gastrocnemius, ** MD values for erector spinae are from -210 to -240 N

proximal joint, the highest impact on joint moments. This is due to the fact, that joint moments are affected by the masses of distal segments. Hence, the highest discrepancies are observed for hip joint moments. The error propagation in proximal direction can be observed. Direction of propagation can result from the recursive method for the inverse dynamic problem. In adopted approach, equations of motion are solved ranging from the most distal segment [11]. For this reason, the greatest differences are observed for residual forces and moments. Their non-zero values indicate a dynamic inconsistency of model, which arises from incompatibility of model assumptions and simplification and experimental data.

Despite the modest effect of perturbing segment mass and inertia tensor on the net joint moments, the greatest effect on selected muscle forces can be observed for both static optimization and CMC algorithm. All modifications consisted in the segment mass reduction, which resulted in muscle forces reduction. Although the extent depends on the chosen method.

Effectiveness of the static optimization approach depends on the model and the choice of muscle. Muscle forces estimated using this method are prone to fluctuations. However muscles forces obtained for initial model fluctuate around forces obtained

for modified model. The greatest discrepancy in muscle forces can be observed for the following muscle: erector spinae, which straightens the back, iliacus and psoas major, which are strong hip flexors. This dependence reflects differences in the hip flexion/extension moment and residual moments. Iliacus and psoas major muscles generate hip joint moments in sagittal plane. While residual moments, virtually acting on the pelvis can turn out in the back muscles. Simultaneously, mass reduction less affect muscle forces acting on distal segments. Therefore, the error propagation in proximal direction can be observed also in muscle forces estimation using static optimization approach. It is attributed to the fact, than static optimization solves the problem of joint moment distribution to muscle forces.

Another trend can be observed for CMC algorithm. In this case, following muscles have proved to be sensitive to mass uncertainties: medial head of gastrocnemius, soleus, tibialis posterior, tibialis anterior and peroneus longus. The first two are the most powerful plantar flexors, same as peroneus longus. While tibialis anterior is acting antagonistically. Tibialis posterior contracts to produce inversion and assists in plantar flexion.

Changes in forces generated by those muscles are associated with the change of foot mass. Therefore the greatest differences are observed for models with foot segment perturbation. Especially significant differences are during stance phase (from 0 to 60 %). However, results of dynamic simulation for tibialis muscles using CMC are especially sensitive to mass change of each segment of lower limb.

It should be noted, that average differences in joint moments not exceeding a few percent (several Nm), produce even 30 % (more than 50N) differences in psoas major and iliacus forces using static optimization technique and similar differences in tibialis muscles using CMC algorithm. Furthermore, considerably reduced forces of selected muscles are connected with reduction antagonistic muscle forces. Force moments generated by such muscles partially cancel each other. This example shows the restriction of dynamic analysis limited to the net joint moments.

5 Conclusion

In general segment mass uncertainty has limited effect on the net joint moments. However, in the case of adopted recursive method for inverse dynamic solution, there is error propagation in proximal direction. Hence the greatest impact was observed for residual moments and forces acting on pelvis, what have not been mentioned in previous papers on sensitivity analysis. Therefore uncertainty in segment mass influence of dynamic consistency of model, what should not be neglected. An alternative to accurate segment mass measurements can be residual reduction algorithms which determines a mass distribution that are more consistent with ground reaction forces. The greatest effect on selected muscle forces can be observed for both static optimization and CMC results but in the different way. Furthermore very big differences can be seen between results obtained using static optimization and CMC

algorithm. Therefore, an accurate estimation of muscles forces and load condition in musculoskeletal system is still posing a challenge.

References

1. Anderson, F., Pandy, M.: A dynamic optimization solution for vertical jumping in three dimensions. *Comput. Methods Biomech. Biomed. Eng.* **2**, 201–231 (1999)
2. Blemker, S., Asakawa, D., Gold, G., Delp, S.: Image-based musculoskeletal modeling: applications, advances, and future opportunities. *J. Magnet. Reson. Imaging* **25**(2), 441–451 (2007)
3. Delp, S., Anderson, F., Arnold, A., Loan, P., Habib, A., John, C., Guendelman, E., Thelen, D.: Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE Trans. Biomed. Eng.* **54**, 1940–1950 (2007)
4. Dempster, W.: Space requirements of the seated operator. Geometrical, kinematic and mechanical aspects of the body with special reference to the limbs. Technical Report (55-159) (AD 87892). Wright Air Development Center. Air Research and Development Command. Wright-Patterson Air Force Base (1955)
5. Erdemir, A., McLean, S., Herzog, W., van den Bogert, A.: Model-based estimation of muscle forces exerted during movements. *Clin. Biomech.* **22**, 131–154 (2007)
6. Fregly, B., Boninger, M., Reinkensmeyer, D.: Personalized neuromusculoskeletal modeling to improve treatment of mobility impairments: a perspective from european research sites. *J. Neuroeng. Rehabil.* **9**(1), 1–18 (2012)
7. Ganley, K.J., Powers, C.M.: Determination of lower extremity anthropometric parameters using dual energy X-ray absorptiometry: The influence on net joint moments during gait. *Clin. Biomech.* **19**(1), 50–56 (2004)
8. John, C., Seth, A., Schwartz, M., Delp, S.: Contributions of muscles to mediolateral ground reaction force over a range of walking speeds. *J. Biomech.* **45**, 2438–2443 (2012)
9. Kiernan, D., Walsh, M., OSullivan, R., OBrien, T., Simms, C.: The influence of estimated body segment parameters on predicted joint kinetics during diplegic cerebral palsy gait. *J. Biomech.* **47**(1), 284–288 (2014)
10. Kingma, I., Toussaint, H.M., De Looze, M.P., Van Dieen, J.H.: Segment inertial parameter evaluation in two anthropometric models by application of a dynamic linked segment model. *J. Biomech.* **29**(5), 693–704 (1996)
11. Kuo, A.: A least-squares estimation approach to improving the precision of inverse dynamics computation. *Trans. ASM* **120**, 148–159 (1998)
12. McConville, J., Clauser, C., Churchill, T.: *Anthropometric Relationships of Body and Body Segment Moments of Inertia*. Anthropology Research Project Inc., Yellow Springs (1980)
13. Nguyen, T., Reynolds, K.: The effect of variability in body segment parameters on joint moment using Monte Carlo simulations. *Gait Posture* **39**(1), 346–353 (2014)
14. Pamies-Vila, R., Font-Llagunes, J., Cuadrado, J., Alonso, F.: Analysis of different uncertainties in the inverse dynamic analysis of human gait. *Mech. Mach Theory* **58**, 153–164 (2012)
15. Pearsall, D., Costigan, P.: The effect of segment parameter error on gait analysis results. *Gait Posture* **9**(3), 173–183 (1999)
16. Rao, G., Amarantini, G., Berton, E., Favier, D.: Influence of body segments parameters estimation models on inverse dynamics solutions during gait. *J. Biomech.* **39**(8), 1531–1536 (2006)
17. Reinbolt, J., Haftka, R., Chmielewski, T., Fregly, B.: Are patient-specific joint and inertial parameters necessary for accurate inverse dynamics analyses of gait? *IEEE Trans. Bio-med. Eng.* **54**, 782–793 (2007)
18. Riemer, R., Hsiao-Weckler, E., Zhang, X.: Uncertainties in inverse dynamics solutions: a comprehensive analysis and an application to gait. *Gait Posture* **27**(4), 578–588 (2008)
19. Scheys, L., Desloovere, K., Suetens, P., Jonkers, I.: Level of subject-specific detail in musculoskeletal models affects hip moment arm length calculation during gait in pediatric subjects with increased femoral anteversion. *J. Biomech.* **44**(7), 1346–1353 (2011)

20. Silva, M., Jorge, A.: Sensitivity of the results produced by the inverse dynamic analysis of a human stride to perturbed input data. *Gait Posture* **19**(1), 35–49 (2004)
21. Świątek-Najwer, E., Będzinski, R., Jr Dragan, S., Krysztoforski, K.: Investigation of lower limb mechanical axis using 3d sonography and magnetic resonance. *Measurement* **45**, 702–710 (2012)
22. Taddei, F., Martelli, S., Valente, G., Leardini, A., Benedetti, M., Manfrini, M., Viceconti, M.: Femoral loads during gait in a patient with massive skeletal reconstruction. *Clin. Biomech.* **27**(3), 273–280 (2012)
23. Thelen, D., Anderson, F., Delp, S.: Generating dynamic simulations of movement using computed muscle control. *J. Biomech.* **36**, 321–328 (2003)
24. van den Bogert, A., Hupperets, M., Schlarb, H., Krabbe, B.: Predictive musculoskeletal simulation using optimal control: effects of added limb mass on energy cost and kinematics of walking and running. *J. Sports Eng. Technol.* **226**, 123–133 (2012)
25. Wesseling, M., de Groote, F., Jonkers, I.: The effect of perturbing body segment parameters on calculated joint moments and muscle forces during gait. *J. Biomech.* **47**, 596–601 (2014)
26. Żuk, M., Pezowicz, C.: Kinematic analysis of a six-degrees-of-freedom model based on isb recommendation: a repeatability analysis and comparison with conventional gait model. *Appl. Bionics Biomech.* 1–9 (2015)
27. Żuk, M., Świątek-Najwer, E., Pezowicz, C.: Hip joint centre localization: evaluation of formal methods and effects on joint kinematics. In: Pi/etka, E., Kawa, J., Wieclawek, W. (eds) *Information Technologies in Biomedicine*, vol. 4, pp. 56–67. Springer (2014)

Part V
Biomaterials

From the Research on Planning of Chemical Syntheses: Computer-Aided Modeling and Prediction of PGI₂ Metabolic Reactions

Zdzislaw S. Hippe

Abstract Methods of manual design of chemical syntheses of complex molecules are briefly discussed. Some hints on computer-aided planning of syntheses are also given. In experimental part, metabolic transformations of prostacyclin PGI₂ in human body are predicted, using own research tool based on a matrix model of constitutional chemistry, enhanced by machine learning algorithms.

Keywords Matrix model · Synthesis tree · Chemical syntheses planning · Prostacyclin PGI₂

1 Introduction

Planning of chemical syntheses is strictly connected with the modeling (simulation) of reactions. For example, in simulation of VIAGRA hydrolysis in human organism [2] an interesting and surprising model was developed: one of the main metabolites can be the methyl alcohol. It is known that this substance controls and changes distinctly the perception of colors by human eyes. Hence, overseers of pilots of air-planes/helicopters and drivers of busses/trucks/taxis were informed about the necessity of keeping a given time span, between the drug intake and beginning of professional activity. Research into computer-assisted syntheses design (further called CASD), showed also the importance to produce substances having a priori given properties or activity. Both features (properties or activity) should be assigned for many chemicals used now in various areas of science, technology and economy. We are thinking here, say, about additives to fuels, various dyes, corrosion inhibitors, organic catalysts or compounds with a given taste or smell, etc.

Z.S. Hippe (✉)

University of Information Technology and Management, 35-225 Rzeszow, Poland
e-mail: zhippe@wsiz.rzeszow.pl

The paper contains a general description of chemical syntheses realized by chemists manually; then some hints on computer-aided syntheses design are also given. In experimental part metabolic reactions of prostacyclin PGI₂ in human body are predicted, using a matrix model of constitutional chemistry (D-U model) [1], enhanced by machine learning algorithms.

2 Background Theory

Progress in CASD of organic syntheses (both in laboratory or industrial scale) has led to a development of some efficient working standards. The milestone of these standards seems to be a skill to recognize—in a synthesized molecule—important structural fragments, i.e. *substructures* (other terms: *functon* or *synton* are also used [7]). This entity (*substructure*) is a mono- or multi-atomic fragment, identified while looking at the structural formula of the synthesized molecule. Substructure itself can be also a goal of the synthesis. Now, let's consider the following problem: How Chemists Execute Planning of Chemical Syntheses?

2.1 How Chemists Execute Planning of Chemical Syntheses

In synthesis planning chemists use, often intuitively, a great variety of problem-solvers of a large span of finesse, depending on individual capabilities, intellectual power, chemical knowledge available and the scale of difficulty of a structure being synthesized. Attempts at ordering the experience in this domain [1] indicate, that two extreme methodologies can be distinguished. The procedures, fundamentally different with respect to generality and analytical and logical sophistication, are called a *direct-associative approach* and a *logic-centered approach*.

Direct-Associative Approach utilizes abilities of recognizing in a synthesized molecule *substructures* which can be properly “assembled” (in a proper sequence) by means of *known* reactions. A distinct characteristic of this methodology is that it always employs only those reactions which are well known to a chemist. This constitutes one of the basic shortcomings of the method, apart from the others concerning, e.g., association skill. It is hard to estimate what number of individual reactions is employed in practice by an experienced chemist; known is the fact that it is a negligible part of the material available in the literature.

In order to realize our restrictions in volume of memorized chemical reactions, let us recall that in the project SPRESI (development of information database on chemical reactions) about 100,000 reactions are registered yearly! Incidentally, let us tell about extreme difficulties in processing of chemical data, especially data on chemical structures. It results from the number of investigated objects (organic compounds). Actual data says about more than 13 millions compounds (metalorganic compounds,

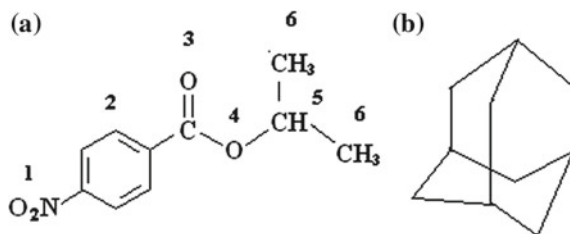


Fig. 1 Illustration of the term “substructure”. Molecule A: nitro group (1), 1,4-disubstituted aromatic ring (2), carbonyl group $>C=O$ (3), C–O–C bond (4), a $-CH<$ fragment (5), and two methyl groups (6). Molecule B: four 6th-membered rings and three 8th-membered rings

polymers and biomolecular substances are not counted). Additionally, phenomenon of structural isomerism complicates this situation. For example, a compound with chemical formula $C_6H_{13}NO_2$ has more than 10,000 isomers, whereas a compound $C_{25}H_{52}$ —above 32 millions of isomers!

This presumably results from two reasons: chemists usually employ a rather restricted stock of their favourite reactions, and—despite experience gained during long work—the number of remembered facts (and hence also reactions) usually undergoes gradual reduction.

Using the classical concept of *functional group*, the following structural fragments may be seen in molecule A (Fig. 1): nitro group, 1,4-disubstituted aromatic ring, carbonyl group $>C=O$, C–O–C bond, a $-CH<$ fragment, and two methyl groups. Yet, the acquired chemical knowledge suggests defining a more complex substructure, namely an ester group $-C(=O)-O-$, as the immediate vicinity of a carbonyl group and a $-O-C-$ bond implies specific chemical properties of this system. Thus, in the case being considered the two separate systems ($>C=O$ and C–O–C) form a common, larger substructure. The borderline of the substructure concept thus becomes diffuse in this case and, accordingly, the kind and number of substructures in a given molecule may vary. Examining again a molecule A we instantly recognize the unit $(CH_3)_2CH-$ as an indivisible structural fragment: it is known that in the process of synthesis, this substructure is transferred intact from the reactant to the product (in other words, it is not synthesized). Interesting questions are dealt with the molecule B (Fig. 1). How many rings are there in this molecule? Careful inspection says about seven rings: four 6th-membered rings and three 8th-membered rings.

Operation of *direct-associative approach* is illustrated by an example of synthesis of compound from cyclopentadiene and methyl vinyl ketone via the Diels-Alder reaction (Fig. 2). Any chemist dealing with the synthesis would find this pathway for the synthesis of compound (III) using a minimum of planning and logical analysis simply because formation of specific substructures occurring in the molecule. Creation of structural fragments present in molecule (III) (a six-membered ring with a double bond, a characteristic bridge) is so obvious and well known that straightforward process of mental association instantly suggests a possible solution. Evidently, this method of synthesis design is confined to relatively simple problems of organic synthesis.

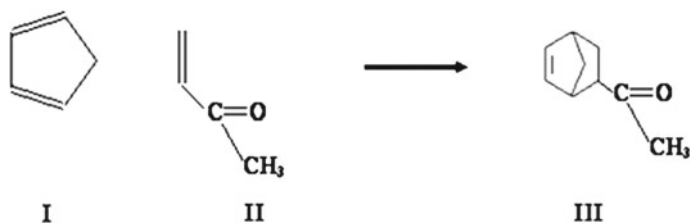


Fig. 2 Application of direct-associative approach in organic synthesis: synthesis of compound (III) from cyclopentadiene (I) and methyl vinyl ketone (II) via the Diels-Alder reaction

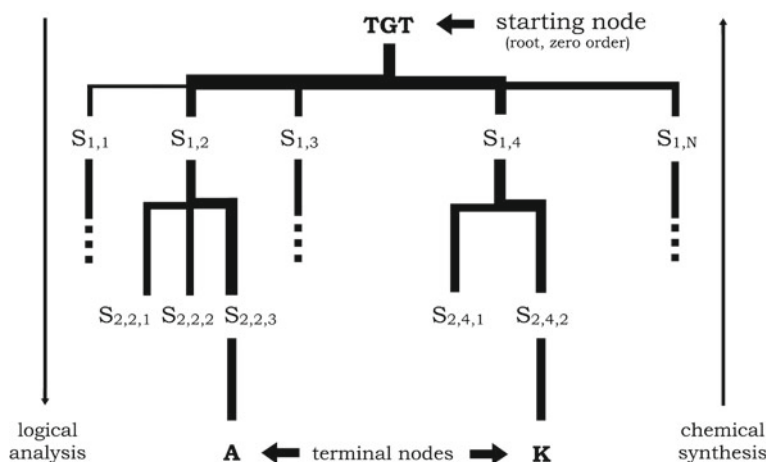


Fig. 3 Synthesis tree (fragment) for molecule TGT, expanded backwards

Logic-Centered Approach is a completely different methodology, restricted only by body of knowledge and time of implementation in the process of synthesis design. The molecule being synthesized, called the *target* structure, (TGT), is subjected to logical analysis, the main goal of which is to predict from which compounds one may obtain a target molecule in single-step, known reactions. These compounds form a set of *subgoals* of the first generation. Repetition of the procedure for each subgoal results in the creation subgoals of the second generation, etc. In this manner a tree-like structure, called a *synthesis tree*, is expanded (Fig. 3).

Particular branches and nodes in a synthesis tree represent chemical conversions and defined subgoals (structures), respectively. The process of expanding of a synthesis tree is continued until the generation of subgoals (subgoal) known to a chemist, or it is terminated automatically when the subgoal is readily available as a reactant for organic syntheses, laboratory or industrial. Generated in this way synthetic paths (reversed path of logical analysis) are simultaneously plans. They require further analysis in order to select the plans with the most favourable characteristics (the smallest number of stages, each of them of the highest possible yield, lack of dan-

gerous or poisonous stages). We see that a general diagram of the possible synthesis routes to a given target molecule (a synthesis tree) can be constructed with lines indicated converting reactions, called *retroreactions* (backward, in comparison to a normal flow of chemical reactions). For the application of retrosynthetic (backward) logic in planning of multistep syntheses of complex molecules and its practical implementation in a system LHASA, Elias J. Corey from Harvard University (Cambridge, USA) was honored by Nobel Prize in chemistry [3]. Chemical syntheses of many important compounds were designed using the then system; e.g. the plant hormone gibberellic acid [4], the total synthesis of which (twenty-step!) culminated one of the most intriguing and salient objectives in the area of organic chemistry. Worthy of notice is the fact, that Corey-chemist—almost at the same time as or perhaps even somewhat earlier than professional computer scientists, invented specific, novel data structure (a tree) and its backward searching method [6].

Comment In practice chemists probably plan organic syntheses using none of the discussed methods in a pure form but rather a hybrid of both approaches. Initially, a partial synthesis tree is generated. When a subgoal of sufficiently simple structure is created, that reminds a chemist of facts known to him/her from organic synthesis, utilization of the direct-associative method takes place. The expansion of a synthesis tree for a complex organic molecule using computers constitutes an alluring possibility of reaching for modern and very rapid methods of synthesis design. In the field of synthesis design, the concept of reversing direction of the process with respect to the normal course of a laboratory experiment, that is, expansion of a synthesis tree beginning from a target structure (“backwards”) is the best illustration of problem reduction by the discussed method. “Backward” reasoning eliminates superfluous transformations, giving simultaneously to the process of synthesis design a particular “driving force”. However, even if a synthesis tree is developed automatically using sophisticated computer program systems, common feature of any synthesis tree is an ability of planning routes of syntheses consisting solely of known (described) reactions; only a sequence of these reactions and/or model for which they were applied can be novel. However, the usage of a matrix model of constitutional chemistry (D-U model) (Sect. 3), is characterized by theoretical possibilities to generate new, i.e. unprecedented chemical reactions. Hence it appears that in the area of simulation of chemical reactions, using a D-U model, it is possible to generate new knowledge. In no other fields of artificial intelligence was it possible to achieve this level of solutions, because intelligent IT systems currently available are only able to reproduce knowledge contained in the knowledgebases, and unable to detect wholly new facts or relations. To summarize hints on automatic planning of chemical syntheses, we may say that reactions generated by intelligent computer systems can be classified into three categories:

1. the reactions whose evaluation, even superficial, indicates full possibility of their laboratory implementation;
2. the reactions which display a considerable degree of ingenuity and lead to interesting conclusions on the strategy of synthesis of the considered compound.

Frequently, these may be the reactions altogether not taken into account, requiring careful checking and evaluation referring to original literature data; and

- the reactions which are (seem to be) unrealistic or too complicated to be carried out in the current chemical knowledge.

3 Experimental. Modeling and Prediction of PGI₂ Metabolic Reactions

The hormone prostacyclin PGI₂ [5] was discovered in 1976 by an international research group in which an important role played Polish scientists R. Gryglewski and A. Szczeklik. For the discovery of prostacyclin and its function in human body, some members of the group (B.I. Samuelson and J.R. Vane) were honored in 1982 by Nobel Prize in physiology. Prediction of PGI₂ metabolic reactions was carried out using a multifunctional CSB system [1], applying the following variants of the matrix model:

- variant D-U (suggested for modeling of new, unprecedented chemical transforms).
Results: 45 reactions; reaction enthalpy -25 to 124 kcal/mol,

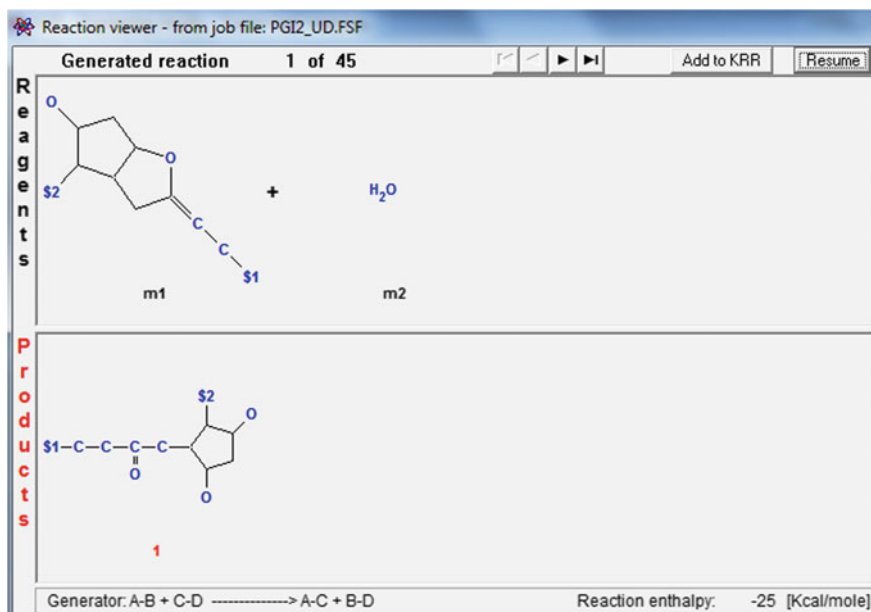


Fig. 4 Possible products of metabolic changes of the hormone in a human body. A co-reagent of PGI₂ is the water molecule. \$1—stands for $-\text{CH}_2-\text{CH}_2-\text{COOH}$, whereas \$2—stands for the fragment $-\text{CH}=\text{CH}-\text{CH}(\text{OH})-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_3$

- variant CS (patterned on a common-sense reasoning of a very gifted chemist during the design of chemical syntheses). Results: 8 reactions; reaction enthalpy -25 to 21 kcal/mol,
- variant SM (similarity model, combined with some machine learning algorithms. Modeling performed in acidic medium). Results: 3 reactions; reaction enthalpy -25 to -19 kcal/mol.

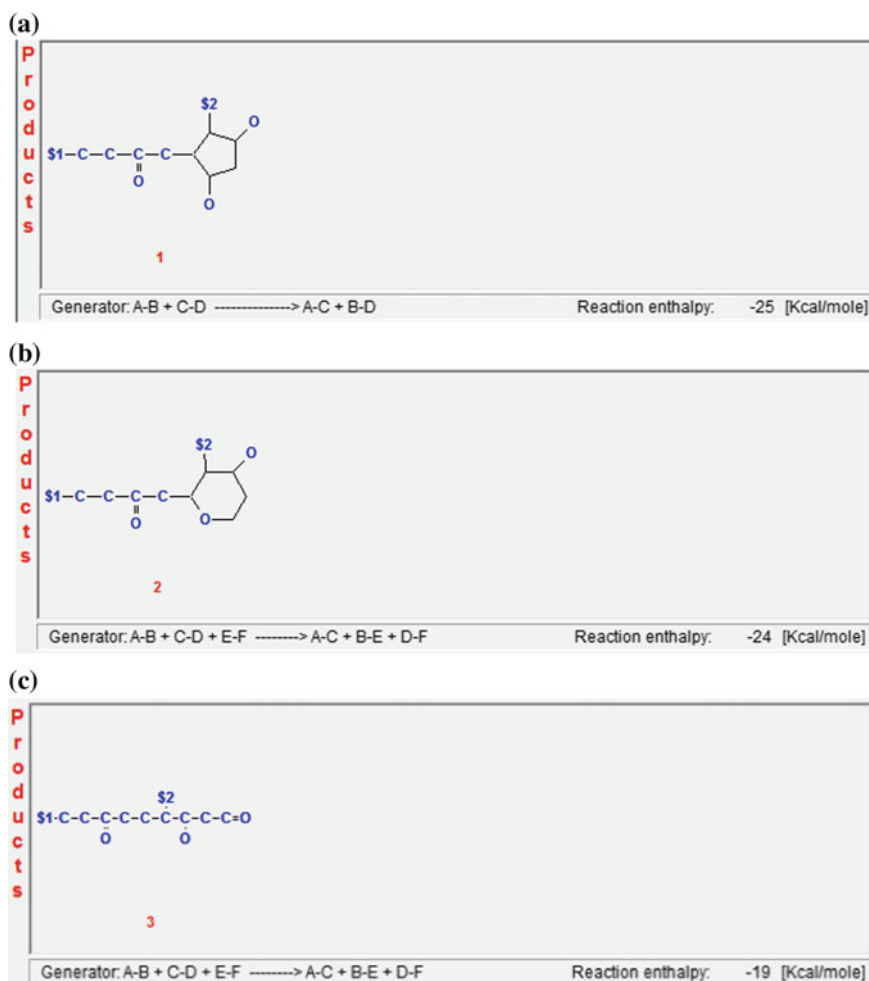


Fig. 5 Prediction of PGI₂ metabolic reactions: variant D-U (products only). **a** Reaction #1 of 45, model D-U, **b** reaction #2 of 45, model D-U, **c** reaction #3 of 45, model D-U, **d** reaction #5 of 45, model D-U, **e** reaction #6 of 45, model D-U, **f** reaction #7 of 45, model D-U

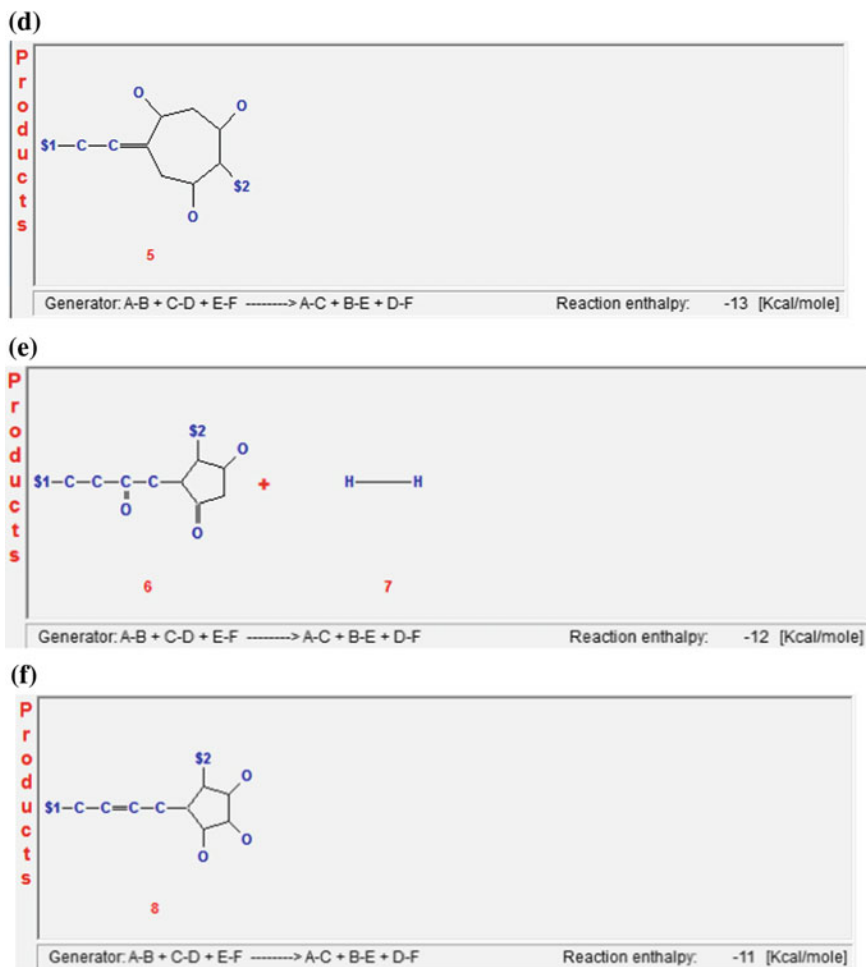


Fig. 5 (continued)

Developed reactions (Fig. 4) disclose possible products of metabolic changes of the hormone in a human body. Therefore, a co-reagent of PGI₂ is the water molecule. The PGI₂ structure is fixed in a simplified way, with accentuation of a central part, which yielded changes denoted as \$1 and \$2 (Fig. 4).

In relation to Reagents all screens shown here have the same content (the same substances: PGI₂ and H₂O molecule. Therefore, in Figs. 5, 6 and 7 only Products are presented. More reactions are available in [1]).

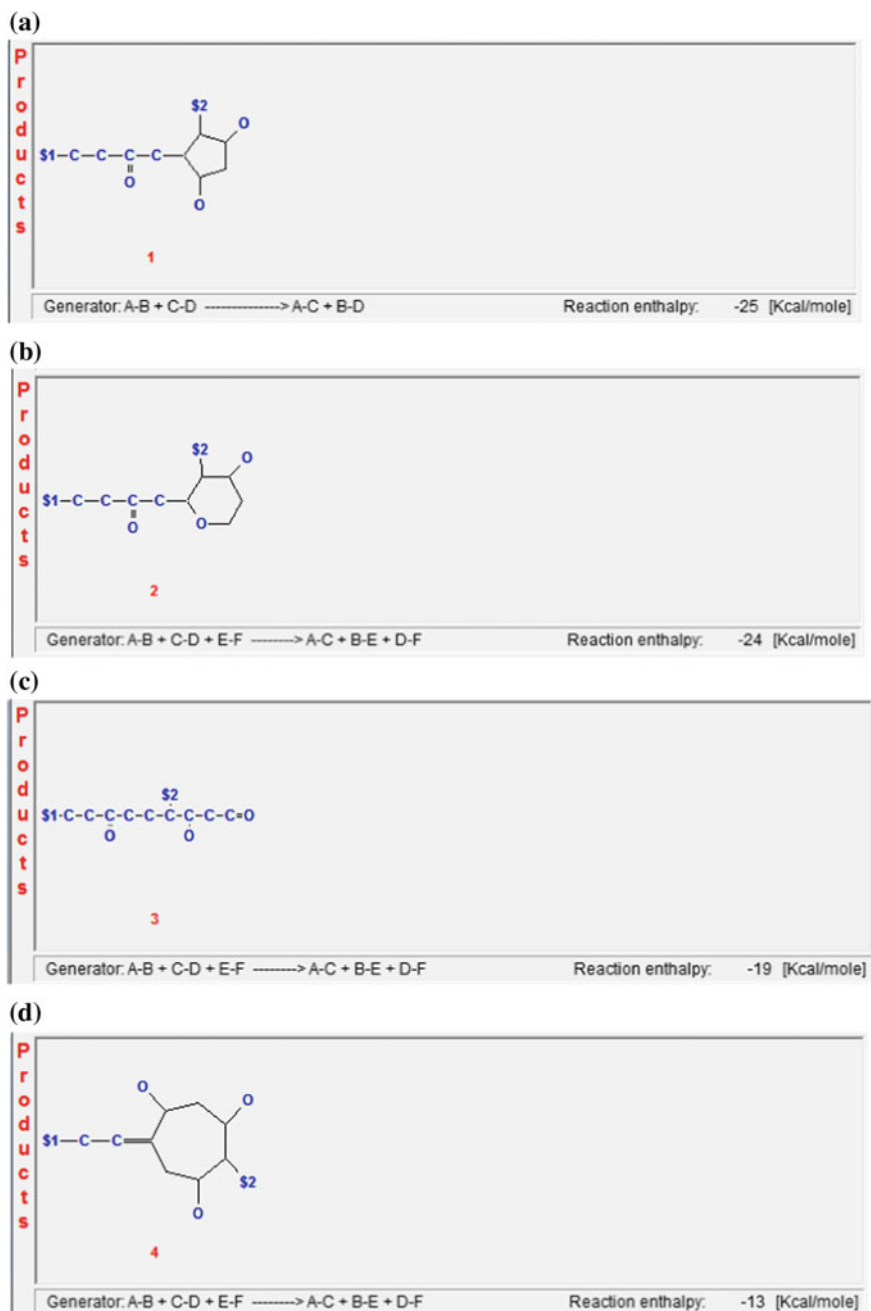


Fig. 6 Prediction of PGI2 metabolic reactions: variant CS (products only). **a** Reaction #1 of 8, model CS, **b** reaction #2 of 8, model CS, **c** reaction #3 of 8, model CS, **d** reaction #4 of 8, model CS, **e** reaction #5 of 8, model CS, **f** reaction #6 of 8, model CS, **g** reaction #7 of 8, model CS, **h** reaction #8 of 8, model CS

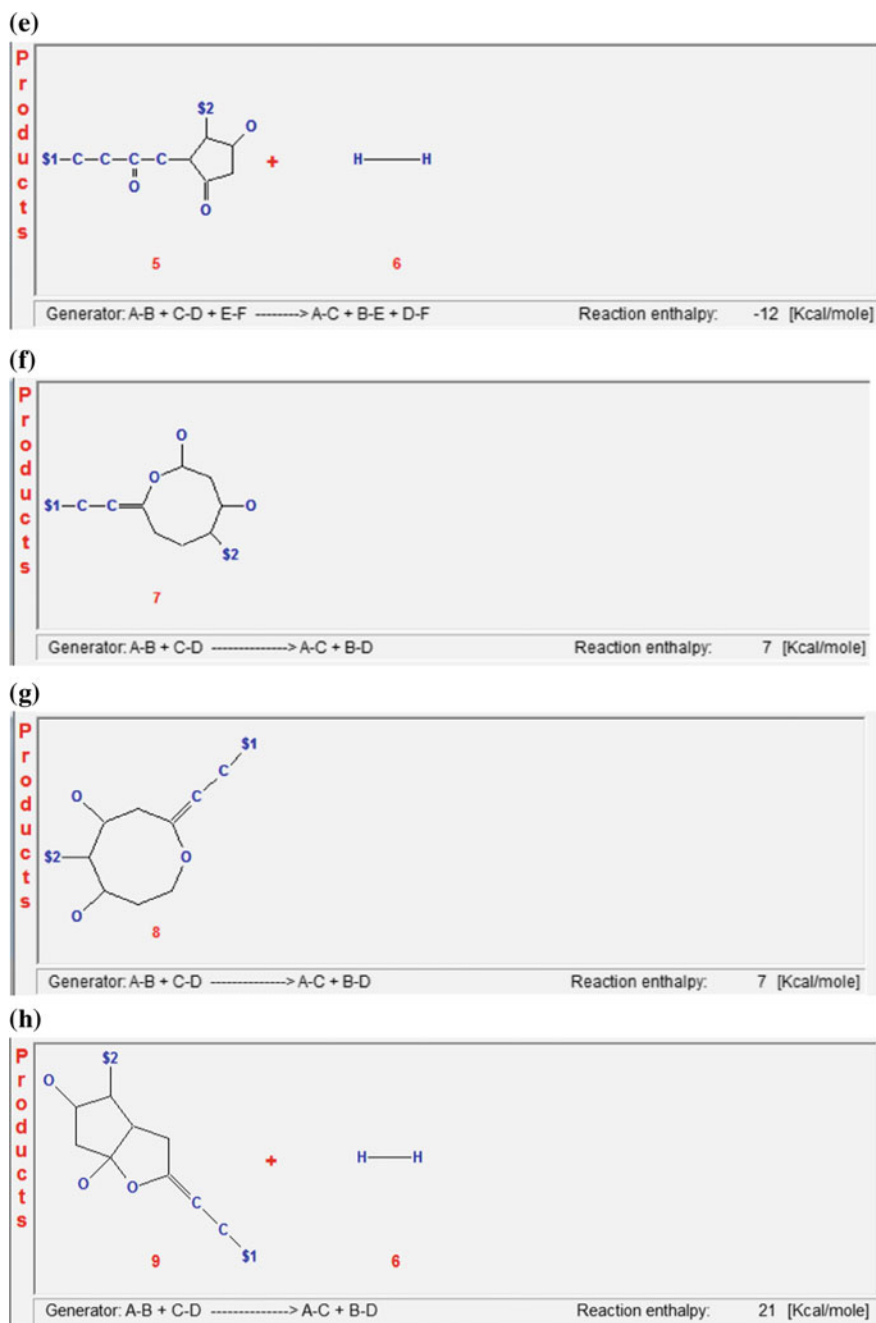


Fig. 6 (continued)

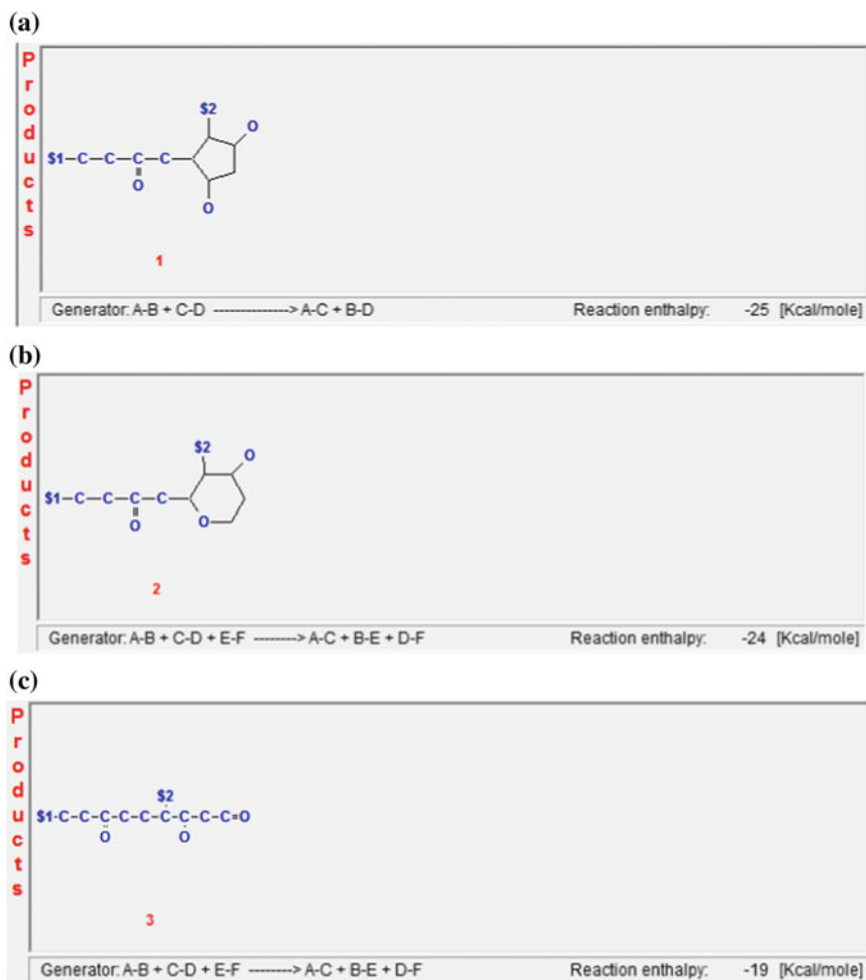


Fig. 7 Prediction of PGI₂ metabolic reactions: Variant SM (acidic medium). Products only. **a** Reaction #1 of 3, model SM, **b** reaction #2 of 3, model SM, **c** reaction #3 of 3, model SM

4 Conclusion

Generated transforms apt reveal the assumed metabolic reactions of PGI₂ hormone in a human body. The first three reactions (for D-U and CS models) have been checked and verified in extensive analytical and clinical investigations by R. Gryglewski. Among PGI₂ metabolites, 95% of the product #1, 4% of product #2 and 0.9% of product #3 were found.

Acknowledgments I would like to express deepest thanks to Prof. R. Gryglewski from Collegium Medicum J. University for generous and independent activity with testing syntetic pathways proposed by the CSB system.

References

1. Grzymala-Busse, J.W., Hippe, Z.S., Mroczek, T.: Exploration of Medical Data—Artificial Intelligence in Planning Syntheses of Drugs. University of Information Technology and Management/Scientific Publisher IVG, Rzeszów-Szczecin (2015) (in Polish)
2. Hippe, Z.S.: A Comparison of Two Approaches Used for Intelligent Planning of Complex Chemical Syntheses. In: Hippe, Z.S., Kulikowski, J.L., Mroczek, T., Wtorek, J. (Eds.) Issues and Challenges in Artificial Intelligence (Studies in Computational Intelligence), vol. 559, pp. 77–88. Springer International Publishing, Switzerland (2014)
3. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1990/corey-lecture.pdf. Accessed 3 Nov 2015
4. https://en.wikipedia.org/wiki/Gibberellic_acid. Accessed 13 Nov 2015
5. <https://en.wikipedia.org/wiki/Prostacyclin>. Accessed 12 Nov 2015
6. Knuth, D: The Art of Computer Programming: Fundamental Algorithms, pp. 308–423. Addison-Wesley (1997) (Section 2.3: Trees)
7. Warren, S., Wyatt, P.: Organic Synthesis: The Disconnection Approach, 2nd edn. Wiley, Chichester (2008)

Electrochemical Properties of TiO₂ Oxide Film on 316LVM Stainless Steel for Orthopedic Implants

Agnieszka Hyla and Witold Walke

Abstract Electrochemical properties of metallic biomaterials in tissue environment have a direct effect on the biocompatibility of implants manufactured out of them. On the other hand, improved corrosion resistance is primarily related to the chemical composition of the surface layer which is in direct contact with bone tissue. The paper proposes modification of the surface of the 316LVM steel with TiO₂ oxide layer using the ALD method, because of the good tolerability of the compound in the human body. In order to determine the optimum number of cycles, and thereby a sufficient thickness of the layer, an electrochemical study (EIS) and corrosion resistance study (potentiostatic and potentiodynamic) were conducted. Studies were carried out in Ringer's solution at the temperature of $T = 37^{\circ}\text{C}$ and $\text{pH} = 6.8 \pm 0.2$. Finding out what is the optimal thickness of the TiO₂ layer has a prospective importance and will contribute to the development of technological conditions with explicit parameters for depositing oxide coatings on implants made out of steel 316LVM. Based on the results it was found that the most favorable electrochemical properties in Ringer solution were shown by TiO₂ layer deposited using ALD method with 500 cycles.

Keywords 316LVM · TiO₂ layer · Electrochemical properties · Ringer's solutions

1 Introduction

Modifying the surface layer increases the biocompatibility of metallic implants. This effect is usually received by forming layers or coatings with electrochemical properties controlled during implants use [20]. 316LVM steel is one of the most frequently used metallic materials for orthopedic implants. Its main advantage are economic aspects. Short-term implants are frequently used in bone surgery. Therefore, the use of implants having a higher biocompatibility, for example titanium alloys, is not economically justified in every case. This is why surface modifications by applying

A. Hyla (✉) · W. Walke

Faculty of Biomedical Engineering, Silesian University of Technology, Zabrze, Poland
e-mail: agnieszka.hyla@polsl.pl

different kinds of layers (mainly oxide) improving steel's biocompatibility and consequently corrosion resistance are conducted more and more often. Thus, the authors proposed the application of TiO_2 layers on steel 316LVM using the ALD method, which is currently one of the most frequently used methods due to the immutability of geometrical characteristics of the modified implant. The main factor influencing implant biocompatibility are its electrochemical properties. That's why the authors focused on them in the article. Taking under consideration the complexity and kinetics of processes concerning bone stabilizers, they need to be characterized by high pitting and crevice corrosion resistance. Fretting corrosion is also to be taken under consideration when it comes to corrosion resistance, it is not the subject of this study though. Possible damage to the protective layer in the environment of the human body can not be a source of electrochemical corrosion initiation and should soon be subject to repassivation. The results of the corrosion resistance of AISI 316LVM steels in vitro environment are discussed as an initial assessment conditioning their suitability for use in medicine. Main component of the passive layer of this steel is Cr_2O_3 oxide. Limited protective properties of this passive layer and the presence of iron and nickel resulted in research on the possible modifications. Therefore striving to improve the biocompatibility of AISI 316LVM steels fully justifies modification of their surface layer. Titanium, carbon or silicon coatings can be considered as layers improving biocompatibility. These actions are aimed, on one hand, at the preparation of modified surfaces for subsequent surface treatments, and on the other hand at providing adequate adhesion, proliferation and cell migration. Due to superhydrophobic effect [6], photocatalytic properties and good biocompatibility the most thermodynamically stable titanium of the n-type semiconductor non-stoichiometric structure of TiO_2 is taken under consideration in such studies. This oxide is distinguished by a wide passivation range reaching even a few volts [10, 19]. Due to the optical properties of TiO_2 formed oxide layers also exhibit different colors depending on the thickness of the layer [14, 28], which may be useful for identification of particular products in the range of sizes.

Literature data indicates [9, 15, 30] that the passive layer, the main component of which is TiO_2 oxide, provides biocompatibility and reduces the penetration of ions of the metallic substrate to the surrounding tissues and body fluids. Many methods are used to modify the surface of TiO_2 , especially electrochemical [5, 18, 22]. In recent years, oxide layers deposition using sol-gel processes and ALD method are becoming more and more frequent object of interest of researchers [4, 16, 25]. One of the most important aspects in regard to the selection of the surface modification method is to ensure the stability of geometrical characteristics over the entire length, particularly of implants in miniaturized form, for example microplates, staples or mini screws. Therefore, in this context, the most suitable surface modification is the application of a layer using the ALD method (atomic layer deposition) [1, 7, 17]. Based on previous research conducted by the authors of this study [2], focused on mechanical properties test and adhesion to the substrate, a beneficial effect on physicochemical properties of 316LVM steel by the application of the TiO_2 layer using ALD method was found. It has been found that the most favorable set of physicochemical properties is present in case of layers applied at a temperature of

approx. 200 °C. In addition to the process temperature, a very important parameter in determining the properties of the layer is the numbers of cycles which affects the layer thickness. Therefore, the authors have attempted conducting an analyzes of selected physicochemical properties of TiO₂ layer deposited by ALD method at the surface of 316LVM steel with a variable number of cycles in the range from 500 up to 2500. The number of cycles of the ALD process has a significant impact on the final quality of the surface layer formed this way, regardless of the type of substrate.

2 Methodology

During the study 316LVM steel samples obtained from hardened rod, diameter of $d = 14$ mm, were used. Samples were subjected to preliminary surface modification consisting of the vibratory treatment using suitable ceramic bodies, and then depositing a layer of TiO₂ by Atomic Layer Deposition (ALD) at different process parameters. The influence of the number of cycles on electrochemical properties of the obtained TiO₂ layers was analyzed. It is assumed that one cycle corresponds with the thickness of the TiO₂ layer of 0.2 Å [32]. A variable number of cycles of 500 (10 nm)–1250 (25 nm) and 2500 (50 nm) and constant temperature for all variations of $T = 200$ °C was used. TiO₂ films studied in this work were grown from TiCl₄ and H₂O in a flow type low-pressure ALD reactor. The deposition process consisted of repeated ALD cycles. Each cycle included a TiCl₄ pulse, purge time, H₂O pulse followed by subsequent.

Pitting Corrosion Test Tests were performed according to the ASTM G48-11 standard [27]. A measuring system consisted of a VoltaLab PGP201 potentiostat, the reference electrode (saturated calomel electrode SCE KP-113), the auxiliary electrode (platinum electrode PTP-201), an anode (the test sample) and a PC with VoltaMaster 4 software. Corrosion tests were started by open circuit potential E_{OCP} . Polarization curves were recorded from the value of initial potential $E_{init} = E_{OCP} - 100$ mV. The potential value changed towards the anode at a polarization rate 0.167 mV/s. After obtaining the corrosion current density of $i_{corr} = 1$ mA/cm² the direction of polarization was changed. Based on these curves corrosion potential E_{corr} and potential of transpassivation E_{tr} were determined. Using the Stern method polarization resistance value R_p was determined. The tests were carried out in Ringer's solution manufactured by Baxter at the temperature of $T = 37 \pm 1$ °C, $pH = 7 \pm 0.2$ (Fig. 1).

Crevice Corrosion Test Tests of resistance to crevice corrosion were performed using potentiostatic method, by recording the change in current density at the potential of $E = +800$ mV at $t = 15$ min. The study used the same measurement system as in the case of the potentiodynamic study.

EIS Test In order to obtain additional information on electrochemical properties of the surface of the analyzed samples, Electrochemical Impedance Spectroscopy (EIS) was conducted. This method allows the analysis and interpretation of processes

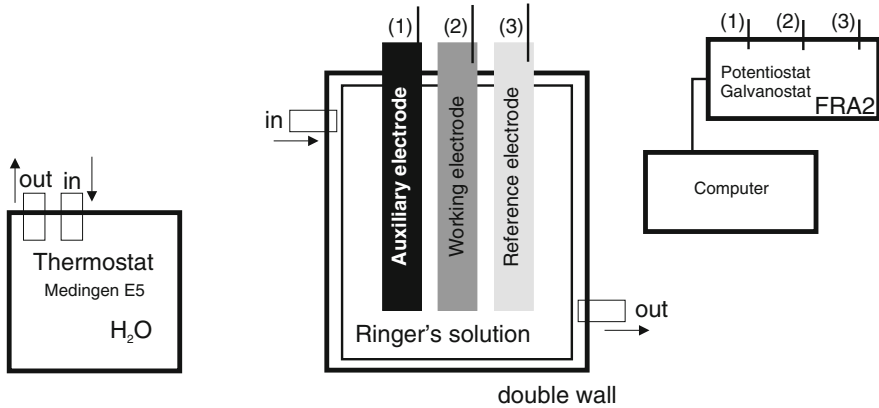


Fig. 1 Scheme of the corrosion test

and phenomena occurring at the interface: implant—tissue environment. The measurements were carried out using the AutoLab PGSTAT302N measurement system equipped with a FRA 2 (Frequency Response Analysis) module. The studies used the same electrode arrangement as in the potentiodynamic studies (Fig. 1). The results provided information about leaks and defects in the surface layers and barrier properties with respect to Ringer's solution. The study determined the impedance spectrum of the system. Obtained data were fitted to the equivalent electrical circuit using least squares method. Impedance spectra of the test are shown in the form of a Nyquist plots for different frequencies (10^4 – 10^{-3} Hz) and Bode diagrams. The amplitude of the sinusoidal voltage of excitation signal is 10 mV [31].

3 Results

The first phase of the work included the study of resistance to pitting corrosion. Tests on samples coated with of TiO₂ layer (500 cycles) showed that the average value of the corrosion potential was $E_{corr} = -147$ mV. The polarization of tested samples in the positive direction enabled determination of the values of transpassivation potential which in this case was equal $E_{tr} = +1421$ mV (Fig. 2). Designated additionally average value of polarization resistance and corrosion current density were equal: $R_p = 551$ k Ω cm² and $i_{corr} = 0.047$ μ A/cm² respectively. The process of increasing the number of cycles influenced the characteristics of the corrosion resistance of TiO₂ layer. Decrease of the average value of corrosion potential to $E_{corr} = -320$ mV (1250 cycles) and $E_{corr} = -361$ mV (2500 cycles) was recorded. The polarization of tested samples also caused an increase of passivation current density at lower values of potentials of transpassivation ($E_{tr} = +1320$ mV and $E_{tr} = +1370$ mV): Fig. 2 and Table 1. Designated average values of polarization resistance and anode

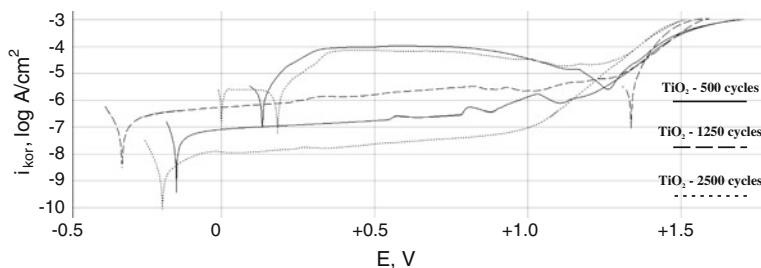
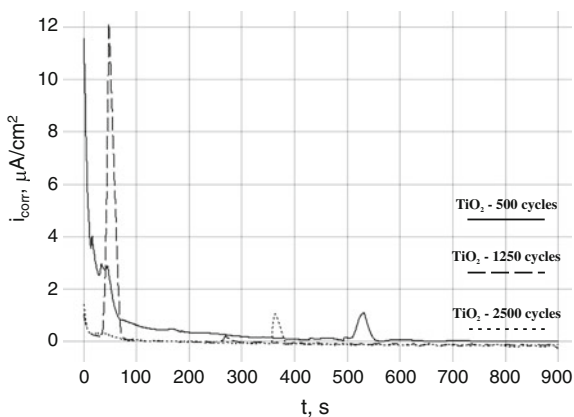


Fig. 2 Typical polarization curves for 316LVM steel samples

Table 1 Results of pitting corrosion resistance

TiO ₂ surface modification	E_{corr} (mV)	E_{tr} (mV)	R_p ($k\Omega cm^2$)	i_{corr} ($\mu A/cm^2$)
500 cycles, 200 °C	-147	+1421	551	0.047
1250 cycles, 200 °C	-320	+1320	460	0.056
2500 cycles, 200 °C	-191	+1370	213	0.122

Fig. 3 Anodic current density as a function of time for 316LVM samples of surface modified by TiO₂ layer



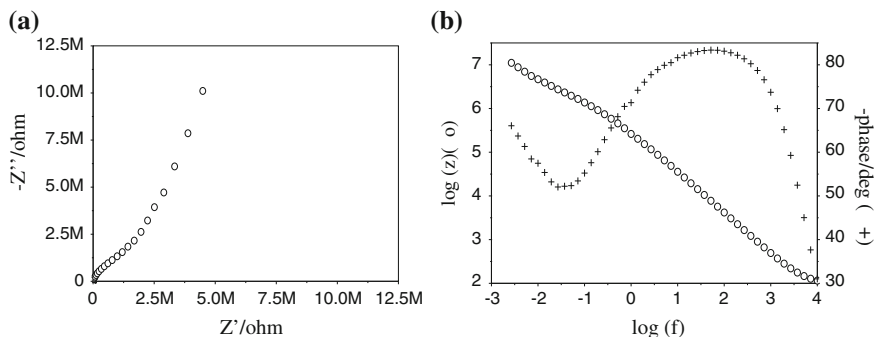
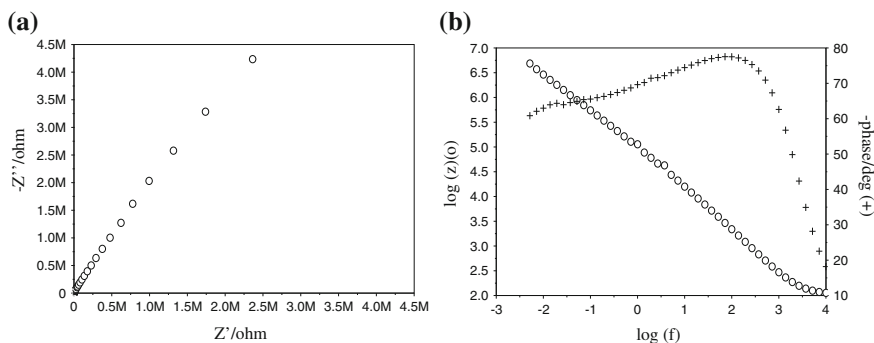
current density were equal as follows: $R_p = 213 k\Omega cm^2$ and $i_{corr} = 0.122 \mu A/cm^2$ —1250 cycles; $R_p = 460 k\Omega cm^2$ and $i_{corr} = 0.056 \mu A/cm^2$ —1250 cycles respectively (Table 1).

The results of changes in anode current density in a function of time in the test of resistance to crevice corrosion show, that regardless of the number of cycles, TiO₂ layer obtained on 316LVM steel provides resistance to crevice corrosion (Fig. 3 and Table 2).

For a fuller assessment of the impact of number of cycles on electrochemical properties of TiO₂ layer in the last stage of the study the EIS test was conducted. During the study impedance spectra were recorded for the hardened 316LVM steel for different numbers of cycles applying the TiO₂ layer (Figs. 4, 5 and 6).

Table 2 Results of crevice corrosion resistance

TiO ₂ Surface modification	Resistance to crevice corrosion	i_{corr} (20 s) ($\mu\text{A}/\text{cm}^2$)	i_{corr} (900s) ($\mu\text{A}/\text{cm}^2$)	Electric charge density (mC/cm^2)
500 cycles, 200 °C	Yes	3.11	0.01	0.399
1250 cycles, 200 °C	Yes	0.13	0.01	0.131
2500 cycles, 200 °C	Yes	0.29	0.01	0.048

**Fig. 4** Impedance spectra for TiO₂ layer (500 cycles): **a** Nyquist plot, **b** Bode diagram**Fig. 5** Impedance spectra for TiO₂ layer (1250 cycles): **a** Nyquist plot, **b** Bode diagram

Determined impedance spectra indicate the diverse kinetics of corrosion processes occurring in the AISI 316LVM—Ringer's solution system. To optimize the equivalent electrical circuit reflecting the process of corrosion in the studied systems, the authors used optimization procedure using the Levenberg-Marquardt method, which is based on a numerical algorithm, minimum mean square error.

The use of two constant-phase elements in an equivalent electrical circuit was proposed, which positively influenced the quality of the fit for curves determined experimentally. R_s is the resistance of Ringer's solution, CPE_{pore} and a constant-phase element used to model the capacity of surface zone of the material with high

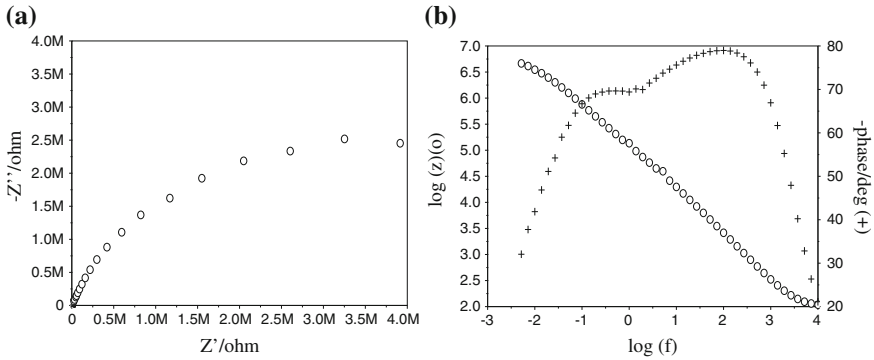


Fig. 6 Impedance spectra for TiO₂ layer (2500 cycles): **a** Nyquist plot, **b** Bode diagram

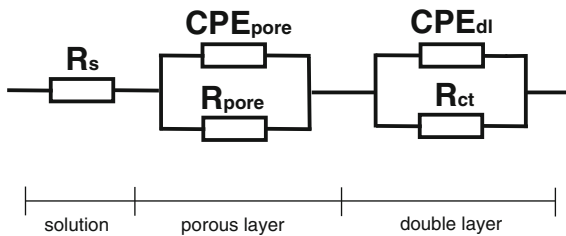


Fig. 7 Electric equivalent circuit and result impedance of corrosive circuits of surface layers formed on the 316LVM steel [8, 29, 31]

Table 3 Results of EIS

TiO ₂ surface modification	R _s (Ωcm ²)	R _{pore} (kΩcm ²)	CPE _{pore}		R _{ct}	CPE _{dl}		E _{OCP}
			Y ₀	n		Y ₀	n	
500 cycles, 200 °C	26	1413	0.8999E-6	0.92	56.7	0.8855E-5	0.89	-185
500 cycles, 200 °C	27	131	0.3674E-5	0.89	13.9	0.3338E-5	0.86	-194
500 cycles, 200 °C	25	38	0.2834E-5	0.97	5.6	0.2002E-5	0.87	-211

surface development, R_{pore}—an element reflecting the resistance of the electrolyte in the zone, on the other hand R_{ct} and CPE_{dl} are resistance and volume of the double layer (Fig. 7, and the mathematical model presented in Eq. (1)).

Based on the patterns of substitution characteristic values describing the corrosion resistance were determined (Table 3).

$$Z = R_s + Z_{eq} \tag{1}$$

$$Z_{eq} = \left(\frac{1}{R_{pore}} + Y_0(j\omega)^n \right)^{-1} + \left(\frac{1}{R_{ct}} + Y_0(j\omega)^n \right)^{-1}$$

4 Conclusion

Corrosion of metallic biomaterials leads to loss of their structural integrity and surface functionality. Metal ions penetrating into human tissue environment may cause an allergic or oncogenic reaction. Therefore, understanding and accurate assessment of corrosion behavior of metallic biomaterials in human tissue environment still remains a subject of interest to many researchers [11–13, 21]. 316LVM steel has a very low corrosion rate in the tissue environment. It has been classified as a passive alloy. Its low corrosion rate is due to the formation of a Cr_2O_3 oxide film (passive coat layer) formed on the surface. Despite these advantages, this layer does not fully protect the steel against environment of corrosive fluids which are human tissue and body fluids. Therefore, application of various kinds of coatings, including Ti-based layers, is more and more often taken under consideration in order to improve biocompatibility. Ti is an element with much better electrochemical properties compared to Cr. One of currently used methods of surface modification is ALD. This method allows researchers to precisely deposit a geometrically constant layer of proper thickness in a very uniform manner on entire implants surface. Results of previous studies showed that the temperature of the process has an important influence on final properties and quality of the layer [2, 3]. Therefore, the authors assessed the influence of the TiO_2 layer thickness on its electrochemical properties in contact with Ringer's solution simulating the actual conditions used for the implant in bone surgery. In order to evaluate the electrochemical behavior of the TiO_2 layer a potentiostatic, potentiodynamic and impedance studies were conducted. The results of these studies clearly showed that regardless of the thickness, the TiO_2 layer improves the resistance of 316LVM steel to pitting corrosion. The obtained results confirm scientific reports from Marin et al. [23, 24] and Shan et al. [26], who also found the influence of ALD process on increasing corrosion resistance of austenitic steel positive, compared to the material in the initial state. In particular, they noticed decrease of the current density and an increase of the perfect passivation area as compared to samples without the applied oxide layer.

Polarization curves measured in potentiodynamic tests were not characterized by the occurrence of hysteresis loop proving the presence of the initiation and development of pitting process. There was only observed a potential of transpassivation in areas of aggregation of ions adsorbed from the solution, mostly Na^+ and Cl^- . This phenomenon is mainly due to the migration of Cl^- ions, assisted by an electric field to the oxide layer, thereby causing its dissolution. The conducted studies showed, that thickness of the layer doesn't improve its electrochemical stability. It has been shown that the layer produced in 500 cycles is characterized by the most advantageous set of properties protecting steel against Ringer's solution environment. In both studies, potentiodynamic and impedance, this layer has obtained the highest value of polarization resistance and charge transfer resistance which testifies its high electrochemical tightness and durability. Increasing the number of cycles and thereby the thickness of the TiO_2 also resulted in lowering the corrosion potential thereby weakening the barrier properties. Regardless of the number of cycles, based on the

potentiostatic tests it was found that the 316LVM steel with the TiO₂ layer is fully resistant to crevice corrosion.

Summing up, the conducted studies clearly demonstrated that the most favorable electrochemical properties were obtained for the TiO₂ layer (10 nm).

References

1. Aarik, L., Arroval, T., Rammula, R., Mändar, H., Sammelselg, V., Aarik, J.: Atomic layer deposition of TiO₂ from TiCl₄ and O₃. *Thin Solid Films* **542**, 100–107 (2013)
2. Basiaga, M., Staszuk, M., Walke, W., Opilski, Z.: Mechanical properties of ALD TiO₂ layers on stainless steel substrate. *Materialwissenschaft Werkstofftechnik* issue **5**, 35–41 (2016)
3. Basiaga, M., Jendruš, R., Paszenda, Z., Kaczmarek, M., Popczyk, M.: Influence of surface modification on properties of stainless steel used for implants. *Arch. Metall. Mater.* **60**(4), 2955–2959 (2015)
4. Basiaga, M., Paszenda, Z., Karasiński, P., Szewczenko, J.: The effects of a SiO₂ coating on the corrosion parameters cpTi and Ti-6Al-7Nb alloy. *Biomater* **4**(1), e28535 (2014). doi:10.4161/biom.28535
5. Basiaga, M., Paszenda, Z., Walke, W., Karasiński, P., Marciniak, J.: Electrochemical impedance spectroscopy and corrosion resistance of SiO₂ coated cpTi and Ti-6Al-7Nb alloy. *Inf. Technol. Biomed. Adv. Intell. Syst. Comput.* **284**, 411–420 (2014). Springer
6. Carp, O., Huismas, C.L., Reller, A.: Photoinduced reactivity of titanium dioxide. *Prog. Solid State Chem.* **32**, 33–177 (2004)
7. Díaza, B., Swiatowska, J., Maurice, V., Seyeux, A., Normand, B., Härkönen, E., Ritala, M., Marcus, P.: Electrochemical and time-of-flight secondary ion mass spectrometry analysis of ultra-thin metal oxide (Al₂O₃ and Ta₂O₅) coatings deposited by atomic layer deposition on stainless steel. *Electrochim. Acta* **56**, 10516–10523 (2011)
8. Fekry, A.M., El-Sherif, R.M.: Electrochemical corrosion behavior of magnesium and titanium alloys in simulated body fluid. *Electrochim. Acta* **54**, 7280–7285 (2009)
9. Hannawa, T., Asami, K., Asoka, K.: Repassivation of titanium and surface oxide film regenerated in simulated bioliquid. *J. Biomed. Mater. Res.* **40**, 530–538 (1998)
10. Houmarda, M., Nunesb, E.H.M., Vasconcelosb, D.C.L., Berthomé, G., Joudc, J.-C., Langletd, M., Vasconcelosba, W.L.: Correlation between sol-gel reactivity and wettability of silica films deposited on stainless steel. *Appl. Surf. Sci.* **289**, 218–222 (2014)
11. Kaczmarek, M.: Investigation of pitting and crevice corrosion resistance of NiTi alloy by means of electrochemical methods. *Electr. Rev.* **86**(12), 102–105 (2010)
12. Kajzer, A., Kajzer, W., Dzielicki, J., Matejczyk, D.: The study of physicochemical properties of stabilizing plates removed from the body after treatment of pectus excavatum. *Acta Bioeng. Biomech.* **2**, 35–44 (2015)
13. Kajzer, A., Kajzer, W., Semenowicz, J., Mroczyk, A.: Corrosion resistance of hip endoprosthesis cups in the initial state and after implantation. *Solid State Phenom.* **227**, 523–526 (2015)
14. Karambakhsh, A., Afshar, A., Malekinejad, P.: Corrosion resistance and color properties of anodized Ti-6Al-4V. *J. Mater. Eng. Perform.* **21**, 121–127 (2012)
15. Kelly, E.J.: Electrochemical behavior of titanium. *Mod. Aspect Electrochem.* **14**, 319–324 (1988)
16. Leskela, M., Ritala, M.: *Angewandte Chemie International Edition* **42**, 5548 (2003)
17. Lim, B.S., Rahtu, A., Gordon, R.G.: Atomic layer deposition of transition metals. *Nat. Mater.* **2**, 749–754 (2003)
18. Liua, X., Chub, P., Dinga, Ch.: Surface modification of titanium, titanium alloys, and related materials for biomedical applications. *Mater. Sci. Eng. R* **47**, 49–121 (2004)
19. Liu, X., Chu, P.K., Ding, C.: Surface modification of titanium, titanium alloys and related materials for biomedical applications. *Mater. Sci. Eng.* **R47**, 49–121 (2004)

20. Long, M., Rack, H.J.: Titanium alloy in total joint replacement—a science perspective. *Biomaterials* **19**, 1621–1639 (1998)
21. Lukaszczuk, J., Śmiga-Matuszowicz, M., Jaszcz, K., Kaczmarek, M.: Characterization of new biodegradable bone cement compositions based on functional polysuccinates and methacrylic anhydride. *J. Biomater. Sci. Polym. Ed.* **18**(7), 825–842 (2007)
22. Marciniak, J., Szewczenko, J., Kajzer, W.: Surface modification of implants for bone surgery. *Arch. Metall. Mater.* **60**(3B), 13–19 (2015)
23. Marin, E., Guzman, L., Lanzutti, A., Ensinger, W., Fedrizzi, L.: Ultralayer Al₂O₃/TiO₂ atomic layer deposition coatings for the corrosion protection of stainless steel. *Thin Solid Films* **522**, 283–288 (2012)
24. Martin, E., Lanzutti, A., Paussa, L., Guzman, L., Fedrizzi, L.: Long term performance of atomic layer deposition coatings for corrosion protection of stainless steel. *Mater. Corros.* **66**(9), 909–914 (2015)
25. Saleem, M.R., Silfsten, P., Honkanen, S., Turunen, J.: *Thin Solid Films* **520**, 5442–5446 (2012)
26. Shan, C.X., Hou, X., Choy, K.-L.: Corrosion resistance of TiO₂ films grown on stainless steel by atomic layer deposition. *Surf. Coat. Technol.* **202**, 2399–2402 (2008)
27. Standard: ASTM G48—11 (2015) An ASTM designation number identifies a unique version of an ASTM standard. G48—11 (2015) G = corrosion, deterioration, and degradation of materials; 48 = assigned sequential number 11 = year of original adoption (or, in the case of revision, the year of last revision) (2015) = year of last reappraisal: Standard Test Methods for Pitting and Crevice Corrosion Resistance of Stainless Steels and Related Alloys by Use of Ferric Chloride Solution
28. Szewczenko, J., Jaglarz, J., Basiaga, M., Kurzyk, J., Skoczek, E., Paszenda, Z.: Topography and thickness of passive layers on anodically oxidized Ti6Al4V alloys. *Electr. Rev.* **88**(12B), 228–231 (2012)
29. Tamilselvi, S., Raman, V., Rajendran, N.: Corrosion behaviour of Ti-6Al-7Nb and Ti-6Al-4V ELI alloys in the simulated body fluid solution by electrochemical impedance spectroscopy. *Electrochim. Acta* **52**, 839–846 (2006)
30. Textor, M., Sitting, C., Frauchiger, V., Tosatti, S.: Properties and Biological Significance of Natural Oxide Films on Titanium and its Alloys. *Titanium in Medicine*. Springer, Heidelberg (2001)
31. Walke, W., Paszenda, Z., Basiaga, M., Karasiński, P., Kaczmarek, M.: EIS study of SiO₂ Oxide Film on 316L stainless steel for cardiac implants. *Inf. Technol. Biomed. Adv. Intell. Syst. Comput.* **284**, 403–410 (2014). Springer
32. Zhang, L., Prosser, J.H., Feng, G., Lee, D.: Mechanical properties of atomic layer deposition-reinforced nanoparticle thin films. *Nanoscale* **4**, 6543–6552 (2012)

Evaluation of Physicochemical Properties of TiO₂ Layer on AISI 316LVM Steel Intended for Urology

Wojciech Kajzer, Anita Kajzer, Magdalena Grygiel-Pradelok,
Anna Ziębowicz and Bogusław Ziębowicz

Abstract The paper presents results of physicochemical properties of TiO₂ layers deposited in the ALD process on AISI 316LVM steel intended for applications in urology. The steel was subjected to electrolytic polishing and chemical passivation and was coated with a oxide layer of TiO₂ in a low temperature ALD process with predetermined variable number of cycles and a constant temperature T = 200 °C of the process. Were analysed processes of 500, 1250 and 2500 of cycles. The paper shows the results of macroscopic evaluation, study of wettability and surface roughness, potentiodynamic test of pitting corrosion resistance and electrochemical impedance spectroscopy EIS in a solution of artificial urine. Based on these results usefulness of the proposed TiO₂ layers on the AISI 316LVM stainless steel was assessed.

Keywords Artificial urine · Corrosion resistance · EIS · Wettability · Stainless steel · 316LVM · TiO₂ layers · ALD

1 Introduction

Stenosis of ureter or urethra often leads to urinary retention in both organs. This condition poses an immediate threat to health and may cause permanent damage to certain organs [34]. The main reasons for stenosis the ureters and urethra includes acquired defects, prostatic hyperplasia, cancer of the ureter and urethra, and narrowing as a consequence of the operation on the urinary tract. To a lesser extent, arising stenosis are the result of congenital defects [8]. One of the basic method of treatment is the use urological stents. Continuous development of techniques are

W. Kajzer (✉) · A. Kajzer · M. Grygiel-Pradelok · A. Ziębowicz
Faculty of Biomedical Engineering, Department of Biomaterials and Medical Devices
Engineering, Silesian University of Technology, Roosevelta 40, Zabrze, Poland
e-mail: wojciech.kajzer@polsl.pl

B. Ziębowicz
Faculty of Mechanical Engineering, Institute of Engineering Materials
and Biomaterials, Silesian University of Technology,
Konarskiego 18a, Gliwice, Poland

© Springer International Publishing Switzerland 2016
E. Piętka et al. (eds.), *Information Technologies in Medicine*,
Advances in Intelligent Systems and Computing 472,
DOI 10.1007/978-3-319-39904-1_34

increasingly demanding for materials used for this type of implants. This applies particularly to their physical and chemical properties, which should be tailored to the urine environment, in which urological implants are staying. In particular, it is required that they are not subject to corrosion, encrustation, migration, they should be resistant to radiation, therapeutic substances, cracking, bending, stretching, and also displayed the ability to elastic deformations. Still unsolved, a very important issue is the problem encrusting the surface of these implants. It involves a partial or complete overlap of the surface of the crystals of calcium phosphate in the form of hydroxyapatite and magnesium ammonium phosphate. The gradual overgrowing side holes and central canal of the implant leads to obstruction and may stop the flow of urine. Encrustation is observed with various intensity depending on the individual characteristics of the patient, regardless of the implantation time and is caused by infection with bacteria which produces urease, involved in the hydrolysis of urea to ammonia and carbon dioxide, causing an increase pH of urine [13, 21, 24]. In many studies on the problem of encrustation was observed that urinary tract infection in the presence of the implant increases the deposition process of these crystals [15].

Degradation of the implanted biomaterial takes place in an environment of human body fluid predominantly from the operational and mechanical loads and corrosive action of environment [14, 23]. Corrosion resistance is dependent on the type of biomaterial metal and condition, morphology and the crystal structure of the surface layer of metal. Therefore, improvement of functional properties of metal implants used in urology, but also in other fields of medicine, such as interventional cardiology or orthopaedics and traumatology, can be obtained by mechanical and electrochemical processing and modify the surface layer of biomaterials. For this purpose, the oxide layer is prepared on surface, e.g. SiO_2 with suitable structure and physicochemical properties ensure high biocompatibility [3–5, 7, 28, 30, 31]. Besides modification of the itself surface of the metal biomaterial can be used a coating made of polymeric biomaterials or biodegradable polymers which release substances. The purpose of these layers is preventing the development of inflammatory conditions caused by bacteria or fungi, reducing the risk of encrustation. Additionally, for long-term implants, biodegradable polymer coating may assist in the healing process by enabling the epithelial tissue coverage of the stent. The substances released in the degradation process of the polymer coating may be: particles of silver and titanium, ciprofloxacin, triclosan, gentian violet, fluconazole, amphotericin or caspofungin [2, 10, 11, 22, 26, 29, 33]. At the same time, the metal structure of the implant will provide lasting treatment effect narrowed sections of the ureter or urethra [17].

Metal biomaterials used for urinary stents are titanium alloys, including popular shape memory alloys, as well as cobalt alloys and widely used for stents in circulatory system AISI 316LVM stainless steel. Comprehensive solution to improve their functional properties in an environment of urine should be focused on shaping the structure of the surface layer. Biocompatibility is essential, so mechanical properties, electrochemical and surface wettability (mainly connected with topography and chemical structure) directly influenced by it. Biocompatibility studies of AISI 316LVM steel indicate a need to eliminate nickel from the surface and replace it

with other elements which increased biocompatibility in the tissue environment. A significant problem is limiting of urological implant surface encrustation process calcium and magnesium compounds. For this purpose, more often protective layers containing titanium—TiO₂ are used. This oxide is distinguished by a wide passive range, reaching even a value of few volts [27]. It is possible to make this type of layers of good adhesion to the substrate, wear resistant and also protecting steel against corrosion processes. Krishna et al. prepared TiO₂ coatings on stainless steel, and they found that the stainless steel showed improved corrosion resistance with TiO₂ coatings [25]. The resulting oxide layers, due to the optical properties of TiO₂, also show varying colour, depends on the layer thickness [27] which can be useful to identifying particular products in model series on different range of sizes. One of promising methods of applying the TiO₂ layer is surface modification by using ALD (atomic layer deposition) method [1, 9, 18]. The method of “atomic layer deposition” is characterized by the fact, that zone which create a solid phase has been limited to monomolecular adsorption layer. This was achieved by use to alternating cycles—pulses of precursors into the reaction chamber. Moreover, after each pulse of precursor will be removed by products of the reaction and not adsorbed molecules from the reaction chamber, by means of pulse of inert gas. It causes, that in the gas phase precursors do not have contact with each other outside adsorbent layer.

Taking into account problems which are relating to the use of metal biomaterials in urology and improving their functional properties, the aim of this study was to define the influence of number of cycles ALD process on selected physicochemical properties of TiO₂ layer applied to AISI 316LVM steel surface, in an artificial urine. In order to achieve our goal, potentiodynamic corrosion resistance, electrochemical impedance spectroscopy, and wettability and surface topography studies were conducted.

2 Materials and Methods

Austenitic stainless steel AISI 316LVM in the form of discs of diameter $d = 14$ mm and thickness $g = 2$ mm was used in the study. The steel of the chemical composition shown in Table 1.

Before surface modification all the samples were electrochemically polished in the phosphate sulphate acid bath and passivated in 40% HNO₃ (samples in initial state). These treatments are required as basic operations for the surface treatment

Table 1 Chemical composition of the analyzed stainless steel (316LVM)

Element, %mass	C	Si	Mn	P	S	Cr	Mo	Ni
ISO 5832-1:2007	≤0.03	≤1	≤2	≤0.025	≤0.01	17–19	2.25–3	13–15
Analyzed steel	0.022	0.59	1.67	0.015	0.001	17.5	2.75	14.25

of austenitic steels used for implants. Then the prepared samples were coated with a TiO₂ layer with the use of Atomic Layer Deposition (ALD). A variable number of cycles: 500, 1250 and 2500 was used along with the constant temperature $T = 200\text{ }^{\circ}\text{C}$, for all the variants. The TiO₂ films investigated in this study were grown from TiCl₄ and H₂O in a low-pressure ALD reactor [16]. The deposition process consisted of repeated ALD cycles. Each cycle included a TiCl₄ pulse, the purge time, H₂O pulse and another purge time. For evaluation of the effect of TiO₂ film deposition on the physicochemical properties of the assumed surface modification, the authors proposed the potentiodynamic and impedance test, surface topography as well as wettability tests.

2.1 Potentiodynamic Test

The tests were carried out as recommended by ISO 10993–15 standard. The test stand comprised of the VoltaLab PGP201 potentiostat, the reference electrode (saturated calomel electrode SCE), the auxiliary electrode (platinum wire), the anode (test sample) and a PC with VoltaMaster 4 software. The corrosion tests started with determination of the open circuit potential EOCP during the first 120 min. The polarization curves were recorded starting with the initial potential value, $E_{\text{init}} = \text{EOCP} - 100\text{ mV}$. The potential changed along the anode direction at the rate of 3 mV/s. Once the anodic current density reached the value of 1 mA/cm², the polarization direction was reversed. On the basis of the curves the corrosion potential E_{cor} , the breakdown potential E_b , the repassivation potential E_{cp} , were determined along with the value of the polarization resistance R_p , calculated with the use of Stern method.

2.2 EIS Test

In order to achieve additional information on electrochemical properties of the surface of the evaluated samples, electrochemical impedance spectroscopy (EIS) was used. Such method allowed for analysis and interpretation of processes and phenomena occurring at the phase boundary: implant—artificial urine environment. The measurements were made with AutoLab PGSTAT 302N system with FRA 2 (Frequency Response Analyse) module. The system of electrodes applied was identical to that used in potentiodynamic tests. Tests making use of EIS brought some information on tightness and faults of the surface layers as well as barrier properties related to artificial urine. Impedance spectra of the system were determined and the data obtained referred to the equivalent circuit. On such basis the values of resistance R and capacity C of the evaluated systems were presented. The impedance spectra of the assessed system were presented as Nyquist diagrams for different frequency values ($10^4 \div 10^{-3}\text{ Hz}$) and also as Bode diagrams. The amplitude of the sinusoid voltage

Table 2 Chemical composition of artificial urine [20]

Chemical composition	NaCl	KCl	NaHPO ₄	Na ₂ SO ₄	MgSO ₄	C ₆ H ₅ Na ₃ O ₇
Concentration (g/dm ³)	6.17	4.75	4.35	2.41	0.99	0.94

inducing signal was 10 mV. The obtained EIS spectra were interpreted following adjustment to the equivalent circuit with the use of the smallest square method. All electrochemical tests were carried out in the environment of 250 ml artificial urine (Table 2) at the temperature $T = 37 \pm 1$ °C, and $\text{pH} = 5.5 \pm 0.2$.

2.3 Wettability Test

In order to determine the surface wettability of the selected samples, the wetting angle and surface free energy (SFE) were evaluated with the use of Owens–Wendt method. The wettability angle measurements were performed with two liquids: distilled water (θ_w) (by Poch S.A.) and diiodomethane (by Merck). Measurements with a drop of liquid and diiodomethane spread over the sample surface were carried out at room temperature ($T = 23$ °C) at the test stand incorporating SURFTENS UNIVERSAL goniometer by OEG and a PC with Surftens 4.5 software to assess the recorded drop image. 10 drops of distilled water and diiodomethane each, 1.5 μl volume, were placed on the surface of each of the samples. The measurement started 20 s after the drops were dripped. Duration of a single measurement was 60 s at the sampling frequency 1 Hz. The mean values of the wetting angle θ_{av} and the surface free energy γ_s were illustrated graphically. The values of surface free energy (SFE) assumed for the calculations, including their polar and dispersion components were respectively, for distilled water: $\gamma_s^p = 51.0$ and $\gamma_s^d = 21.8$ mJ/m² and for diiodomethane: $\gamma_s^p = 6.7$ and $\gamma_s^d = 44.1$ mJ/m².

2.4 Microscopic Observation and Surface Roughness Measurements

Evaluation of the surface using a stereomicroscope SteREO Discovery V8 from Zeiss with software AxioVision was carried out. While roughness measurement using contact profilometer Surtronic 3+ by Taylor Hobson was performed. The parameter Ra was determined—the arithmetic average of ordinates profile, in accordance with recommendations of PN–EN ISO 4287. Values as average of 5 measurements on each sample with standard deviation were given.

3 Results and Discussion

3.1 Potentiodynamic Test

Results of potentiodynamic tests carried out to evaluate the pitting corrosion resistance are presented in Table 3 and Fig. 1. Results obtained for electrochemically polished and chemically passivated 316LVM stainless steel in initial condition made it possible to define the corrosion potential E_{cor} , the value of which ranged between -242 and -256 mV. It was also observed that the value of transpassivation potential E_{tr} ranged between $+1190$ and $+1207$ mV, while the polarization resistance R_p between 155 and 165 $k\Omega cm^2$. Similar values of the breakdown potential and the polarization resistance were obtained by the authors [14, 23] in investigations of implants manufactured of the stainless steel, after electrochemical polishing.

Subsequently, samples with prepared TiO_2 layer deposited by the ALD process were tested. Regardless of the number of the ALD process cycles, breakdown potential E_b and repassivation potential E_{cp} were observed, demonstrating possibility of development pitting corrosion. Average values of breakdown potential for analysed variations were greater than the transpassivation potential E_{tr} set for initial state and

Table 3 The results of pitting corrosion tests

Surface condition	E_{cor} (mV)	E_b (mV)	E_{cp} (mV)	E_{tr} (mV)	R_p ($k\Omega cm^2$)
Initial state	-249 (10)	—	—	1198 (12)	160 (7)
500 cycles	-206 (5)	$+1630$ (35)	$+1363$ (55)	—	409 (8)
1250 cycles	-172 (5)	$+1625$ (8)	$+1361$ (16)	—	174 (4)
2500 cycles	-175 (6)	$+1659$ (8)	$+135$ (6)	—	184 (14)

E_{cor} —corrosion potential, E_b —breakdown potential, E_{cp} —repassivation potential, E_{tr} —transpassivation potential, R_p —polarization potential

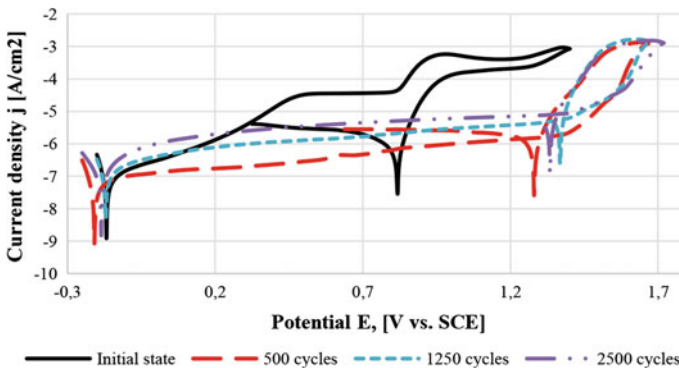


Fig. 1 Polarization curves of AISI 316LVM at initial state and after TiO_2 ALD surface modification with different number of process cycle

were in range from $E_b = +1559$ to $E_b = +1669$ mV. In addition, all registered repassivation potential E_{cp} values were also higher than those established transpassivation potential values for initial state and ranged from $E_{cp} = +1302$ to $E_{cp} = +1380$ mV. This indicates that corrosion resistance was increased of samples subjected to surface modification by titanium oxide in the ALD process compared to initial state of samples in an artificial urine (despite breakdown potentials). For the 500 ALD process cycles increase the value of polarization resistance R_p compared to other analysed options was stated. Furthermore, advantageous reduction in the current density was observed—Table 3 and Fig. 1.

Based on the potentiodynamic tests, taking into account results preferred option was TiO₂ layer applied in 500 ALD process cycles was stated. Similarly, improvement of corrosion resistance of stainless steel and titanium alloys, as a result of applying the layer of titanium dioxide, in the environment simulating human body fluids authors were observed [6, 19, 25].

3.2 EIS Results

In the Figs. 2 and 3 the impedance spectra for AISI 316LVM steel in initial state and with applied TiO₂ layer after process in established technological conditions were shown.

For analysis impedance spectra corrosion system experimentally determined for AISI 316LVM steel + TiO₂—artificial urine, equivalent electrical circuit (as shown in Fig. 4) were used. Obtained impedance spectra for test samples were interpreted by comparison to equivalent electrical circuit, which indicates presence a double layer (two time constants visible in the diagrams), where R_s —resistance of the artificial urine, R_p —electrolyte resistance in pores and CPE_p —capacity of double layer

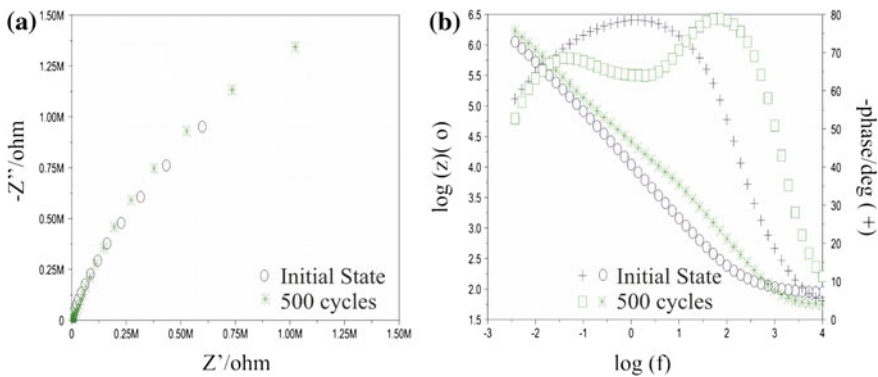


Fig. 2 Impedance spectrums for AISI 316LVM steel in initial state and after 500 cycles ALD process: **a** Nyquist diagram. **b** Bode diagram

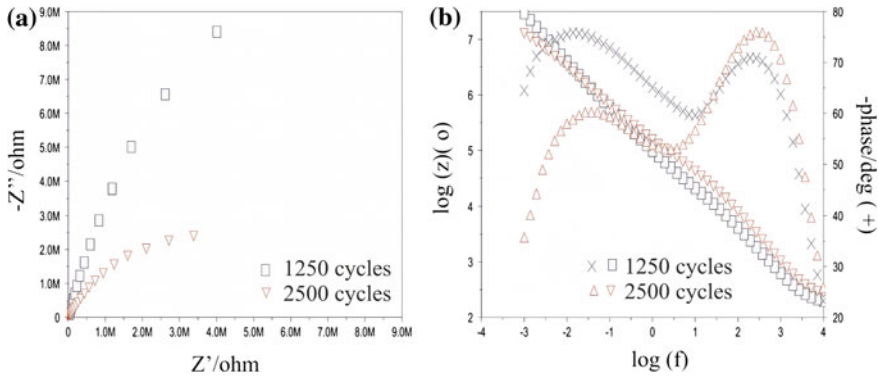


Fig. 3 Impedance spectrums for AISI 316LVM after 1250 and 2500 cycles ALD process: **a** Nyquist diagram. **b** Bode diagram

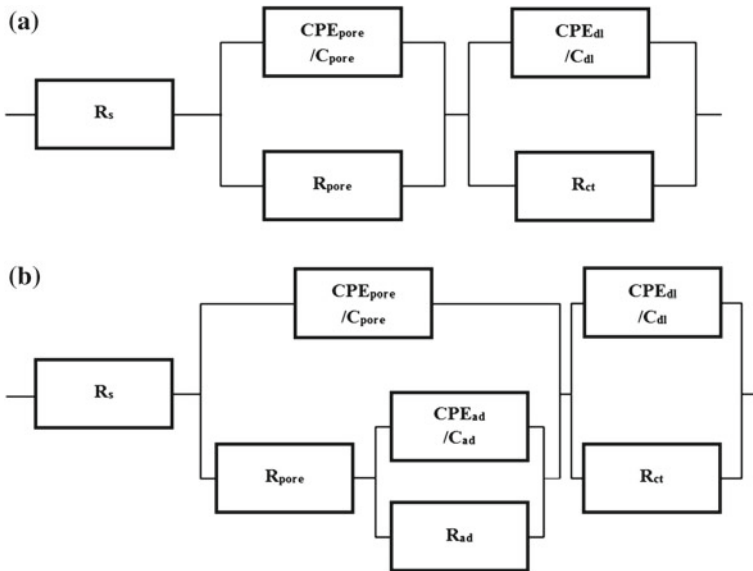


Fig. 4 Equivalent electrical circuit for AISI 316LVM—TiO₂ layer—artificial urine: **a** double layer, **b** triple layer

(porous, outside surface), whereas R_{ct} and CPE_{dl}/C_{dl} —the resistance and capacitance of the oxide layer. Application two elements with the constants phase in equivalent electrical circuit had positively influenced quality of the fit curves experimentally determined for samples in initial state and after 500 cycles ALD process—Fig. 4a. The impedance spectra for samples after 1250 and 2500 cycles ALD process by comparison to equivalent electrical circuit indicating presence of an additional sub-layer were interpreted—Fig. 3b. Suggested fit from processes taking place in porous

Table 4 Results of EIS

Surface	R_s (Ωcm^2)	R_{ad} ($\text{k}\Omega\text{cm}^2$)	C_{ad} (μFcm^{-2})	R_{pore} ($\text{k}\Omega\text{cm}^2$)	CPE_{pore}	
					Y_{01} ($\Omega^{-1}\text{cm}^{-2}\text{s}^{-n}$)	n_1
Initial state	55	—	—	1	0.2621e-4	0.87
500 cycles	48	—	—	63	0.8727e-5	0.84
1250 cycles	51	402	18	9	0.8487e-5	0.87
2500 cycles	56	750	6	21	0.7290e-5	0.76
Surface	C_{dl} (μFcm^{-2})	R_{ct} ($\text{k}\Omega\text{cm}^2$)	CPE_{dl}		E_{OCP} (mV)	
			Y_{02} ($\Omega^{-1}\text{cm}^{-2}\text{s}^{-n}$)	n_2		
Initial state	—	2500	0.1887e-4	0.87	-198	
500 cycles	18	3200	—	—	-151	
1250 cycles	—	4220	0.4793e-5	0.89	-196	
2500 cycles	—	6550	0.1190e-5	0.99	-141	

oxide layer, forming on samples surface by interaction with electrolyte was resulted. Character similar to capacitively porous layer by resistance R_{ad} and capacity C_{ad} adsorption layer were supplemented—Fig. 4b.

In conclusion, the obtained results of impedance measurements prove significant influence of TiO₂ layer on corrosion resistance of AISI 316LVM steel were confirmed—Table 4. It was found, that in all cases impedance modulus tested corrosion systems was reduced with increased frequency. The phase angle also with frequency was changed. Obtained impedance spectra for all samples about high stability of electrochemical oxide layer on such prepared steel surfaces were provided. Additional presence of the porous layer on oxide surfaces layer with highly developed surface area is a barrier against corrosive environment in artificial urine. Successively, in 1250 and 2500 cycles ALD processes formation an additional adsorption sublayer unfavourable processes of initiation encrustation mineral compounds coming from the solution of the artificial urine were evidenced.

3.3 Wettability Results

The results of wettability and surface energy calculations and examples of drops dripped in the surface of samples in initial state and samples with TiO₂ are presented in Figs. 5 and 6.

The obtained results pointed to the increased values of the wetting angle θ_{av} for the samples of stainless steel, type AISI 316LVM with ALD deposited TiO₂ layer. The electrochemically polished and passivated samples are hydrophilic and

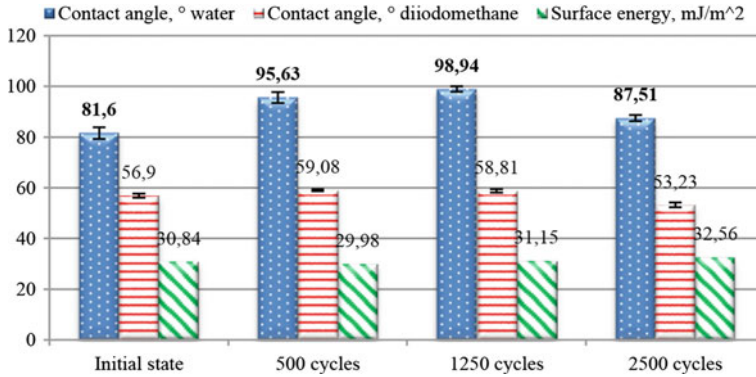


Fig. 5 The surface energy calculated on the basis of the contact angle measurements—OW method

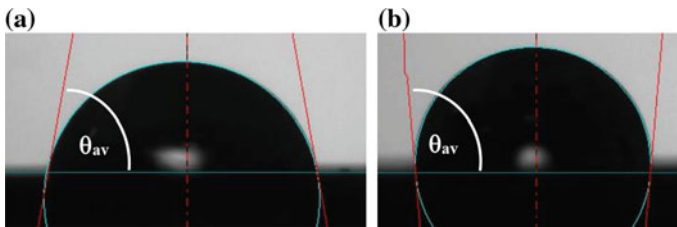


Fig. 6 Sample measurement of wetting angle, sample: **a** initial state ($\theta_{av} = 81.63^\circ$), **b** 500 cycles ($\theta_{av} = 95.63^\circ$)

show very low surface wettability ($\theta_{av} = 81.63^\circ$). Such results were also obtained by the authors [32], who evaluated the physicochemical properties of surfaces of AISI 316LVM steel implants after electrochemical polishing and passivation. The obtained the mean value of the contact angle θ_{av} ranged between 76.44 and 81.95° . Deposition of the TiO_2 film over the studied samples effected in the elevated wetting angle, regardless of the number of cycles in the ALD process, resulting therefore in the hydrophobic surfaces. No significant differences were observed in the values of the surface energy throughout the investigated sample surfaces. The study indicated that steel with the deposited TiO_2 layer showed higher wetting angle ($\theta_{av} = 95.63^\circ$ and $\theta_{av} = 98.94^\circ$) and as compared to the values obtained in surfaces after polishing and passivation. The wetting angle values in samples with the TiO_2 layer (a number of cycles 500 and 1250) were $\theta_{av} > 90^\circ$ which proved the hydrophobic properties of the evaluated surfaces and appeared as a favourable phenomenon for the urology implants [32] contrary to investigations by the authors [12], who grew silicon films on steel samples with the use of sol-gel method, observing increased wettability in all variants of the investigated coatings, as compared to polished samples, which is an adverse feature in case biomaterials intended for urological implants.

3.4 Microscopic Observation and Surface Roughness Measurement

Based on conducted measurements it can be concluded, that TiO₂ layer have no significantly affected of roughness surface change in relation for surface in initial state which was subjected to electrolytic polishing and chemical passivation (Table 5). This is confirmed by the fact, that oxide layer obtained by ALD process provide the constancy of geometrical features of surface. Provide a mapping of output surface condition but not its alignment [6].

Samples surface macroscopic observation (Fig. 7) after potentiodynamic tests relative to surface before testing show presence of corrosion pits (Fig. 7d, f, h). This is confirmed by potentiodynamic curves of samples with TiO₂ layer which breakdown potential and hysteresis loops showing initiation of pitting corrosion and its development were observed (Fig. 1). Only for electropolishing and chemical passivation samples surface where transpassivation potential was observed, after tested, pitting corrosion was not found.

Table 5 Surface roughness

Surface condition	Initial state	500 cycles	1250 cycles	2500 cycles
Ra (μm)	0.14 (0.03)	0.15 (0.04)	0.17 (0.02)	0.16 (0.05)

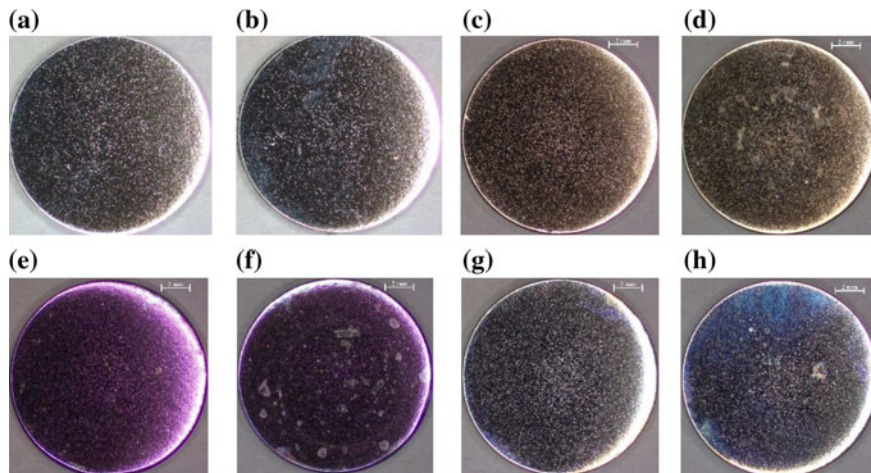


Fig. 7 Sample of surface specimens: **a, b** initial state before and after corrosion test. **c, d** ALD process of 500 cycles before and after corrosion test. **e, f** ALD process of 1250 cycles before and after corrosion test. **g, h** ALD process of 2500 cycles before and after corrosion test—stereoscopic mag. 4.8x

4 Conclusion

Modification of 316LVM steel surface with titanium dioxide is of interest of many researchers who are dealing with issues related to improving biocompatibility of material in the tissue environment. In particular, use TiO_2 on metal biomaterials surface intended for implants used in urology looks promising. An aggressive environment of urine, are leading to early metal biomaterials destruction, and above all, problem with encrustation leads to stopping urine flow, causing necessity for frequent replacement. In view of complicated structure urological stents and their small size, suitable method for application protects oxide layers may be ALD process. This process allows control film thickness during the application process and above all, maintain a constant thickness on whole implant.

Based on potentiodynamic results can be concluded, that suggested variants of surface modification with ALD process, which was carried out at $T = 200^\circ\text{C}$ during while TiO_2 layer was applied, is useful for improvement of physicochemical properties of AISI 316LVM steel in artificial urine. It was found that ALD process of 500 cycles is the most preferred. Obtained the highest value of polarization resistance and the smallest value of the corrosion current compared to other analysed variants of ALD process and also compared to initial state of samples, the smallest surface activity of investigated steel was indicated, Table 1, Fig. 1. Moreover, creating titanium dioxide layer was contributed to increase breakdown potential value and repassivation potential value relative to transpassivation potential for samples having a surface treated with only on electropolishing and passivation process.

On the basis of the EIS results, for the ALD processes of 1250 and 2500 cycles, the presence of additional adsorption sublayer on samples surface was observed. This layer indicates on initialization of encrustation process by mineral compounds coming from the artificial urine solution. This phenomenon is disadvantageous, because canal of vessel slowly will be overgrown reducing the urine flow in consequence. Excessive growth of encrustation will stop the flow completely.

The study of surface topography confirmed that, regardless of the number of ALD process cycles, no change of roughness was observed. Therefore, choosing suitable surface modification which allow to form a predetermined surface roughness before the ALD process is essential. The similar results were obtained in the surface wettability tests. Regardless of the number of cycles, the TiO_2 layer demonstrated comparable surface wettability of hydrophobic character. Such nature of the layer may prevent the implant surface against encrustation in an urological fluid.

In summary, the obtained test results indicate that ALD method is perspective for metal implants surface modification for use in urology. In order to fully characterize TiO_2 layers, further research, in particular as regards to encrustation of urological implant, need to be conducted. The research should include analysis of influence of long-term exposure to artificial urine of AISI 316LVM steel samples modified in the ALD process on their surface encrustation, as well as on their physicochemical properties and mass density of metal ions released to the solution.

References

1. Aarik, L., Arroval, T., Rammula, R., Mandar, H., Sammelselg, V., Aarik, J.: Atomic layer deposition of TiO₂ from TiCl₄ and O₃. *Thin Solid Films* **542**, 100–107 (2013)
2. Bachmann, S.P., Ramage, G., VandeWalle, K., Patterson, T.F., Wickes, B.L., Lopez-Ribot, J.L.: Antifungal combinations against candida albicans biofilms in vitro. *Antimicrob. Agents Chemother.* **47**, 3657–3659 (2003)
3. Basiaga, M., Paszenda, Z., Walke, W., Karasiński, P., Marciniak, J.: Electrochemical impedance spectroscopy and corrosion resistance of SiO₂ coated cpTi and Ti-6Al-7Nb alloy. *Information Technologies in Biomedicine. Advances in Intelligent Systems and Computing*, vol. 284, pp. 411–420. Springer (2014)
4. Basiaga, M., Walke, W., Paszenda, Z., Kajzer, A.: The effect of EO and steam sterilization on mechanical and electrochemical properties of titanium grade 4. *Mater. Tehnol.* **50**(1), 153–158 (2016)
5. Basiaga, M., Walke, W., Paszenda, Z., Karasiński, P., Szewczenko, J.: The effects of a SiO₂ coating on the corrosion parameters cpTi and Ti-6Al-7Nb alloy. *Biomaterials* **4**(1) (2014)
6. Basiaga, M., Walke, W., Staszuk, M.: Study of electrochemical properties of 316LVM steel with TiO₂ layer deposited by means ALD method. In: *International Conference on Advanced Computational Engineering and Experimenting—ACE-X 2015*. Munich, Germany (2015)
7. Basiaga, M., Jendruś, R., Walke, W., Paszenda, Z., Kaczmarek, M., Popczyk, M.: Influence of surface modification on properties of stainless steel used for implants. *Arch. Metall. Mater.* **60**(4), 2965–2969 (2015)
8. Borkowski, A.: *Urologia: Podręcznik dla studentów medycyny*. PZWL, Warszawa (2008). (in Polish)
9. Diaza, B., Swiatowska, J., Mauricea, V., Seyeux, A., Normand, B., Harkonen, E., Ritala, M., Marcus, P.: Electrochemical and time-of-flight secondary ion mass spectrometry analysis of ultra-thin metal oxide (Al₂O₃ and Ta₂O₅) coatings deposited by atomic layer deposition on stainless steel. *Electrochim. Acta.* **56**, 10516–10523 (2013)
10. DiTizio, V., Ferguson, G.W., Mittelman, M.W., Khoury, A.E., Bruce, A.W., DiCosmo, F.: A liposomal hydrogel for the prevention of bacterial adhesion to catheters. *Biomaterials* **19**, 1877–1884 (1998)
11. Hachem, R., Reitzel, R., Borne, A., Jiang, Y., Tinkey, P., Uthamanthil, R., Chandra, J., Ghanoun, M., Raad, I.: Novel antiseptic urinary catheters for prevention of urinary tract infections: correlation of in vivo and in vitro test results. *Antimicrob. Agents Chemother.* **53**, 5145–5149 (2009)
12. Houmarda, M., Nunes, E.H.M., Vasconcelosb, D.C.L., Berthomec, G., Joudc, J.-C., Langletd, M., Vasconcelosba, W.L.: Correlation between sol-gel reactivity and wettability of silica films deposited on stainless steel. *Appl. Surf. Sci.* **289**, 218–223 (2014)
13. Jacobse, S.M., Strickler, D.J., Mobley, H.L., Shirliff, M.E.: Complicated catheter-associated urinary tract infections due to Escherichia coli and Proteus mirabilis. *Clin. Microbiol. Rev.* **21**(1), 26–59 (2008)
14. Kajzer, A., Kajzer, W., Dzielicki, J., Matejczyk, D.: The study of physicochemical properties of stabilizing plates removed from the body after treatment of pectus excavatum. *Acta. Bioeng. Biomech.* **2**, 35–44 (2015)
15. Klimek, L., Różański, W., Jabłonowski, Z., Sosnowski, M., Kliś, R.: Obserwacje zmian mikroskopowych cewników typu Double J w zależności od czasu utrzymywania ich w drogach moczowych. *Inżynieria Biomateriałów* **43–44**, 36–39 (2005). (in Polish)
16. Kumagai, H., Masuda, Y., Shinagawa, T.: Self-limiting nature in atomic-layer epitaxy of rutile thin films from TiCl₄ and H₂O on sapphire (001) substrates. *J. Cryst. Growth* **314**, 146–150 (2011)
17. Lee, G., Marathe, S., Sabbagh, S., Crisp, J.: Thermo-expandable intra-prostatic stent in the treatment of acute urinary retention in elderly patients with significant comorbidities. *Int. Urol. Nephrol.* **37**, 501–504 (2005)

18. Lim, B.S., Rahtu, A., Gordon, R.G.: Atomic layer deposition of transition metals. *Nat. Mater.* **2**, 749–754 (2003)
19. Marciniak, J., Szewczenko, J., Kajzer, W.: Surface modification of implants for bone surgery. *Arch. Metall. Mater.* **60**(3B), 2123–2129 (2015)
20. Opalko, F.J., Adria, J.H., Khanc, S.R.: Heterogeneous nucleation of calcium oxalate trihydrate in artificial urine by constant composition. *J. Cryst. Growth* **181**, 410–417 (1997)
21. Ostrowska, K., Strzelczyk, A., Różalski, A., Strączek, P.: Biofilm bakteryjny jako przyczyna zakażeń układu moczowego—mikroorganizmy patogenne, metody prewencji i eradykacji. *Postepy HigMed Dosw* **67**, 1027–1033 (2013). (in Polish)
22. Roe, D., Karandikar, B., Bonn-Savage, N., Gibbins, B., Rouillet, J.B.: Antimicrobial surface functionalization of plastic catheters by silver nanoparticles. *J. Antimicrob. Chemother.* **61**, 869–876 (2008)
23. Semenowicz, J., Mrocza, A., Kajzer, A., Kajzer, W., Koczy, B., Marciniak, J.: Total hip arthroplasty using cementless advantage cup in patients with risk of hip prosthesis instability. *Ortop. Traumatol. Rehabil.* **16**(3) 07, 253–263 (2014)
24. Shaw, G.L., Choong, S.K., Fry, C.: Ecrustation of biomaterials in the urinary tract. *Urol. Res.* **33**(1), 17–22 (2005)
25. Siva, D., Krishna, R., Sun, Y.: Thermally oxidised rutile-TiO₂ coating on stainless steel for tribological properties and corrosion resistance enhancement. *Appl. Surf. Sci.* **252**, 1107 (2005)
26. Suci, P.A., Mittelman, M.W., Yu, F.P., Geesey, G.G.: Investigation of ciprofloxacin penetration into *Pseudomonas aeruginosa* biofilms. *Antimicrob. Agents Chemother.* **38**, 2125–2133 (1994)
27. Szewczenko, J., Jaglarz, J., Basiaga, M., Kurzyk, J., Skoczek, E., Paszenda, Z.: Topografia i grubość warstw pasywnych na utlenianym anodowo stopie Ti6Al4V. *Przegląd elektrotechniczny* **88**(12B), 228–231 (2012) (in Polish)
28. Szewczenko, J., Jaglarz, J., Basiaga, M., Kurzyk, J., Paszenda, Z.: Optical methods applied in thickness and topography testing of passive layers on implantable titanium alloys. *Opt. Appl.* **43**(1), 173–180 (2013)
29. Trupti, A., Lauserpina, C., Shanta, M.: Efficacy of a silicone urinary catheter impregnated with chlorhexidine and triclosan against colonization with *proteus mirabilis* and other uropathogens. *Infect. Control Hosp. Epidemiol.* **28**, 596–598 (2007)
30. Walke, W., Paszenda, Z., Basiaga, M., Karasiński, P., Kaczmarek, M.: EIS study of SiO₂ oxide film on 316L stainless steel for cardiac implants. *Information Technologies in Biomedicine. Advances in Intelligent Systems and Computing*, vol. 284, pp. 403–410. Springer (2014)
31. Walke, W., Paszenda, Z., Karasiński P., Marciniak, J., Basiaga, M.: Investigations of mechanical properties of SiO₂/TiO₂ coatings deposited by sol–gel method on cpTi and Ti6Al7Nb. In: *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*
32. Xu, L.C.: Effect of surface wettability and contact time on protein adhesion to biomaterial surfaces. *Biomaterials* **28**, 3273–3283 (2007)
33. Yao, Y., Ohko, Y., Sekiguchi, Y., Fujishima, A., Kubota, Y.: Self-sterilization using silicone catheters coated with Ag and TiO₂ nanocomposite thin film. *J. Biomed. Mater. Res. B Appl. Biomater.* **85**, 453–460 (2008)
34. Zabel, M.: *Histologia. Podręcznik dla studentów medycyny i stomatologii*. Elsevier Urban and Partner (2013) (in Polish)

Non-destructive Evaluation of Biomaterial's Magnetic Field After Plastic Deformation

Milan Smetana, Vladimir Chudacik, Klara Capova
and Monika Oravcova

Abstract The article deals with the non-destructive evaluation of conductive biomaterials. Various austenitic steel specimens are inspected under the same conditions. Residual weak magnetic field of the biomaterials is investigated. Commercial fluxgate sensor is used for the inspection. The gained results are presented and discussed in the paper. Biomaterials applied in the form of implants are under long-term mechanical stresses.

Keywords Non-destructive evaluation · Fluxgate sensor · Conductive biomaterials · Austenitic stainless steel

1 Introduction

A role of non-destructive evaluation (NDE) of material structures is undeniable worldwide, because it plays very important role in present time. Increased R&D activities in the field of NDE have been motivated by the need for precise evaluation of cracks and flaws for the assessment of the expected life of mechanical components. Industry still continues to reap the benefits of solid state magnetic field sensing. Moreover, many industrial applications require complex inspection where the main aim is to prevent a failure of various components [1, 6]. Periodic inspection of such components and devices ensures their safe, effective and long-term operation. New methods and devices are still being developed and designed to tackle gradually increasing demands for reliable detection and precise characterization of materials and their defects. NDE techniques are widely used in a range of industries for the

M. Smetana · V. Chudacik (✉) · K. Capova
Faculty of Electrical Engineering, Department of Electromagnetic and Biomedical
Engineering, University of Zilina, Univerzitna 1, 010 26 Zilina, Slovak Republic
e-mail: vladimir.chudacik@fel.uniza.sk

M. Oravcova
Faculty of Mechanical Engineering, Department of Materials Engineering,
University of Zilina, Univerzitna 1, 010 26 Zilina, Slovak Republic

inspection of various complex structures, such like biomaterials [11]. New technologies and activities from the last decades have allowed to design and to develop special types of sensors, which have properties suitable for the electromagnetic NDE applications.

This article deals with non-destructive investigation of conductive biomaterials. Magnetic field of the austenitic stainless steel is inspected before and after applied plastic deformation. The reason for this inspection is to give information about changes of biomaterial's structure. This is based on the residual magnetic field evaluation. Experimental results are presented and discussed in this paper.

2 Conductive Biomaterials and Magnetic Field Evaluation

Biomaterial is defined as a non-living material used in medical environment to interact with living system. Biomaterials are used as prosthesis, implants, medical equipment and instruments intended for internal usage. These materials have to be biocompatible which refers to the properties of biomaterials being biologically compatible with living tissues or a living system without toxicity. The problem of these metallic biomaterials is their relatively large susceptibility to corrosion. High Young's modulus and related strength also sufficient high yield strength are in metallic materials related with toughness and elongation which play predominant role in final properties. These final properties have to be taken into consideration, so the metallic biomaterials could transfer relatively high load without unwanted changes in shape. The final properties depend mostly on their chemical composition and on final processing methods (mechanical processing and heat treatment). Both of the factors influence the structure of the material [8]. Metallic materials have predominant role in orthopedic surgery in most orthopedic devices, including temporary devices (e.g. bone plates, pins and screws) and permanent implants (e.g. total joint replacements). These materials can be categorized according to their major alloying element: stainless steels, cobalt-based alloys and titanium-based alloys [2]. Stainless steel is the generic name for iron-based alloys that contain a high percentage (11–30 %) of chromium and varying amounts of nickel. Stainless steels can be categorized according to chemical composition into two groups: the chromium and chromium-nickel types. Alternatively, they can also be grouped based on the characteristic microstructure of the alloys: martensitic, ferritic, austenitic, or duplex (austenitic plus ferritic). Except for the duplex type, each of the other three groups of stainless steels finds applications in medical devices. However, only austenitic stainless steels are used as implants [8]. The most important group of austenitic stainless steels are chromium-nickel and chromium-nickel-molybdenum types, stabilized with niobium or titanium (e.g. AISI 316Ti), including the types with low carbon content (e.g. AISI 316L) [2]. Conductive biomaterials can be evaluated non-destructively, using special electromagnetic methods. The main aim is to investigate their inhomogeneities (mainly fatigue and stress-corrosion cracks, notches), but also to reveal the structure deformations, due

to the phase transitions [9, 10]. Some special types of conductive biomaterials may change their structure and electromagnetic properties, due to the applied mechanical deformation. Austenitic stainless steels are such types of biomaterials. These materials have the same properties as paramagnetic materials. Mechanical deformation leads to the phase transformations (transitions), where induced martensitic component is generated. Thus, ferromagnetic component is induced. Moreover, austenitic and induced martensitic component have different mechanical properties.

Investigation of the residual (weak) magnetic fields of the conductive biomaterials can be performed using sensitive magnetic field sensors. One of the theoretically well-described is the fluxgate sensor. This magnetometer was firstly introduced in the 1930s. It is a solid-state device for measuring the magnitude of DC, or low frequency AC magnetic fields, in given direction [5]. The sensors are reliable, rugged as they contain no moving parts; they have very low energy consumption and they can reach 10pT resolution and 1nT long-term stability. Fluxgate magnetometers have advantages over other types of field sensors in certain area of field intensities and frequency measurements. Unlike the induction magnetometers, which register field changes only, fluxgate sensors may be used for absolute measurements besides measuring the direction of the field [3]. Even though, good sensitivity is offered by fluxgate sensors with no need for superconductivity [4, 7].

3 Experimental Set-Up

Experimental biomaterial—conductive stainless steel specimens are inspected in this study. Three different austenitic steel grades, according to the AISI standard, are evaluated: 304, 316L and 316Ti grade. The samples have the cylindrical shape, with an initial height (measured from the base to the top) of $h = 14$ mm and with radius of $r = 6$ mm. All the specimens were previously dissolutionally annealed at the defined temperature (1050 °C/35 min.). This regime was defined as the reference state. Further, the controlled mechanical deformation was applied. The deformation was performed by mechanical loading in the opposite sides of the cylindrical specimens, Fig. 1. Four values of the applied plastic deformations were inspected: $T_1 = 5\%$, $T_2 = 10\%$, $T_3 = 20\%$ and $T_4 = 40\%$ (this value represents the percentage shortening of the length of the specimen after the plastic deformation, in comparison with the reference sample).

Every specimen was checked using hardness measurement (according Vickers) to define the deformation percentage. Measurement was performed by measuring the length of the indenter crosslinks. Following inspection is performed: constant lift-off distance is set to $h_1 = 1.6$ cm, above the base of the biomaterial.

The fluxgate magnetometer, shown in Fig. 2 is used to pick-up the response signal. It is the commercial one, made by Canon. Measured component of the magnetic field in given direction is converted to the output voltage signal.

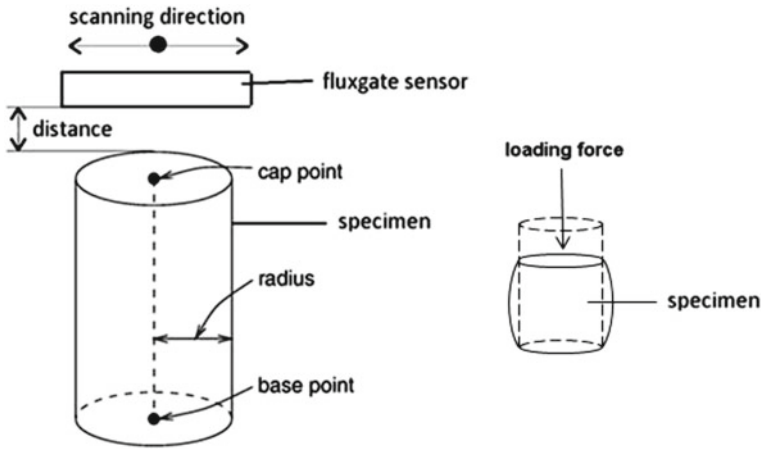


Fig. 1 Spatial configuration of the specimen and probe during the inspection

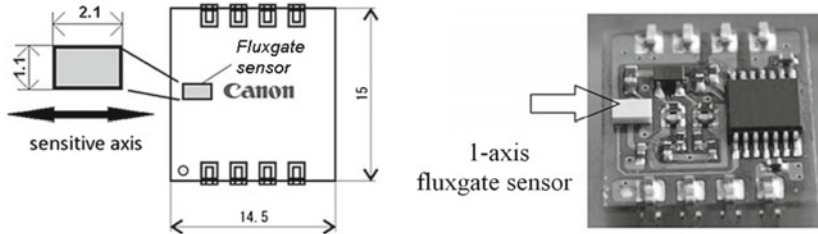


Fig. 2 Dimensions and configuration of the sensing element (fluxgate magnetometer)

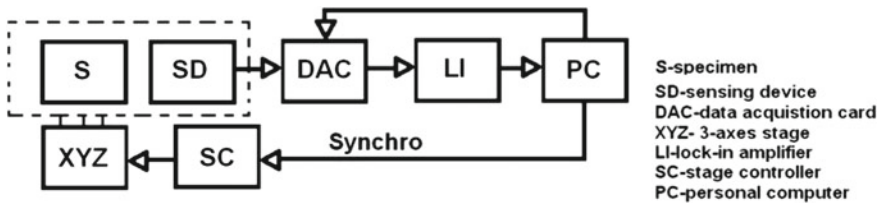


Fig. 3 Flowchart of the experimental set-up

Configuration of the experimental set-up is shown in Fig. 3. Measured data are acquired using the data acquisition card with resolution of $r = 16$ bit/channel, $f_s = 2$ kS/s. User interface for data manipulation, controlling the stage and processing the data is created using the LabVIEW virtual instrumentation. Two dimensional surface raster scan is performed over each sample, in defined rectangular area (number of scanning lines $n = 80$, step distance of $s = 0.25$ mm, scanned length per line of $p = 35$ mm).

4 Experimental Results

Results of the experimental measurements are presented in this section. Three-dimensional waveforms display distribution of the magnetic field over the material's surface. Surface 2D graphs display appropriate maximum values of the magnetic field only. All the waveforms are processed as the differential responses (reference signal is subtracted from the individual waveforms).

The waveforms shown in Figs. 4, 5 and 6 are the illustrative examples. All the waveforms can be displayed in such configuration. It can be seen that the residual magnetic field is distributed on non-uniform way, over the specimen's surface. Moreover, the direction as well as the values of the field is changing. This is the result of the applied loading force, leading to the plastic deformation. To compare all the inspected results, following type of the graphs are presented: only maximum values of the 3D graphs are taken into account, to be the resulting waveforms displayed within one graph.

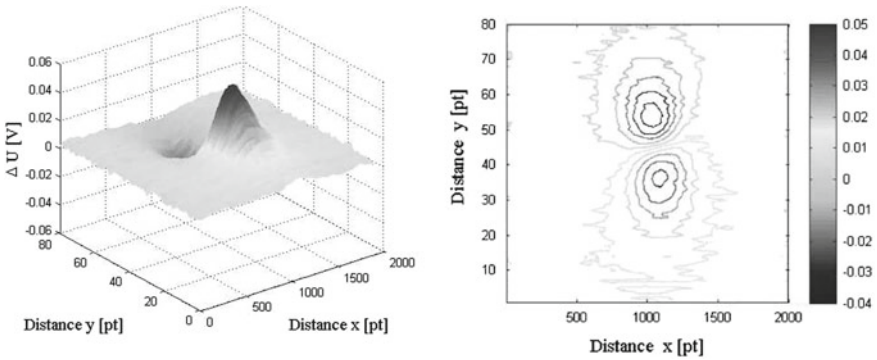


Fig. 4 Experimental results: residual magnetic field distribution, AISI 316Ti, $T_1 = 5\%$

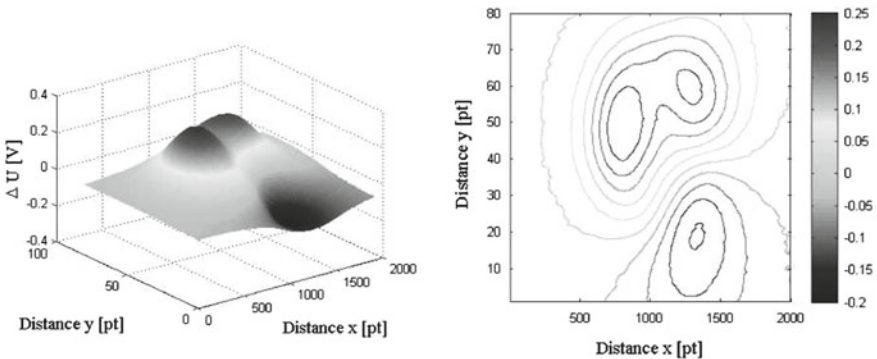


Fig. 5 Experimental results: residual magnetic field distribution, AISI 304, $T_1 = 5\%$

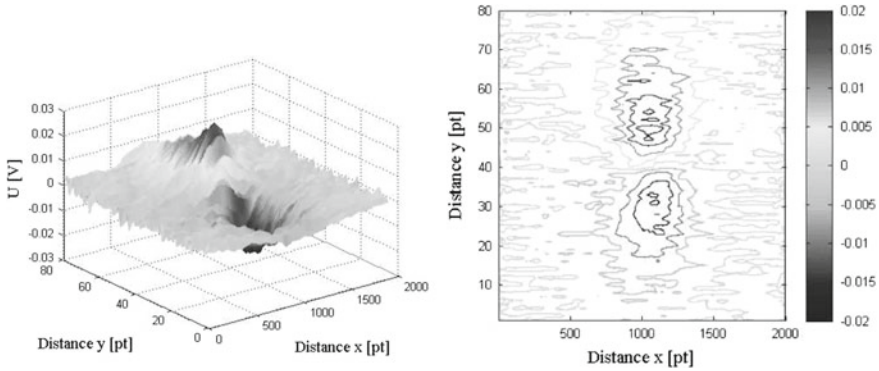


Fig. 6 Experimental results: residual magnetic field distribution, AISI 316L, $T_1 = 5\%$

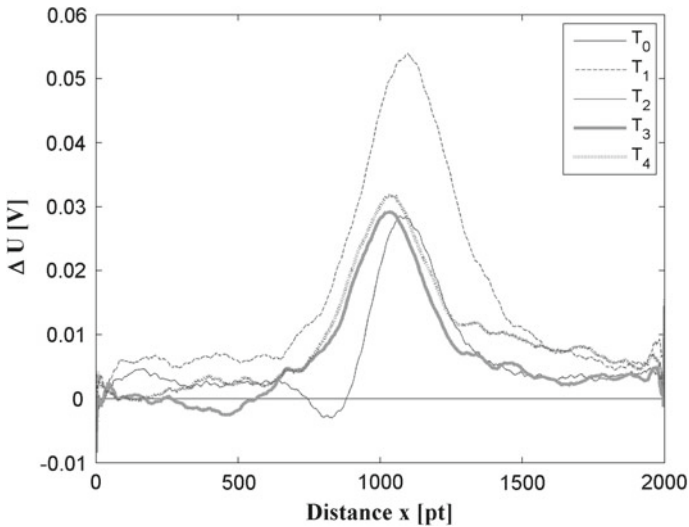


Fig. 7 Experimental results: maximum field values, AISI 316Ti

Figure 7 display the individual waveforms for all the plastic deformation values. It can be observed that the highest response has specimen with the deformation of $T_1 = 5\%$. This lies in fact that the higher the applied deformation, the lower the length of the specimen. The lift-off distance is increasing in such way, what possesses the lower-values response signal.

Individual waveforms presented in Fig 8 present the same situation, for AISI 304 specimens. As can be seen, the highest signal strength is obtained for the plastic deformation of $T_4 = 40\%$. Moreover, all the waveforms are approximately 200 times higher, in comparison with AISI 316Ti results. This is caused by higher level of the induced martensitic component, which is ferromagnetic.

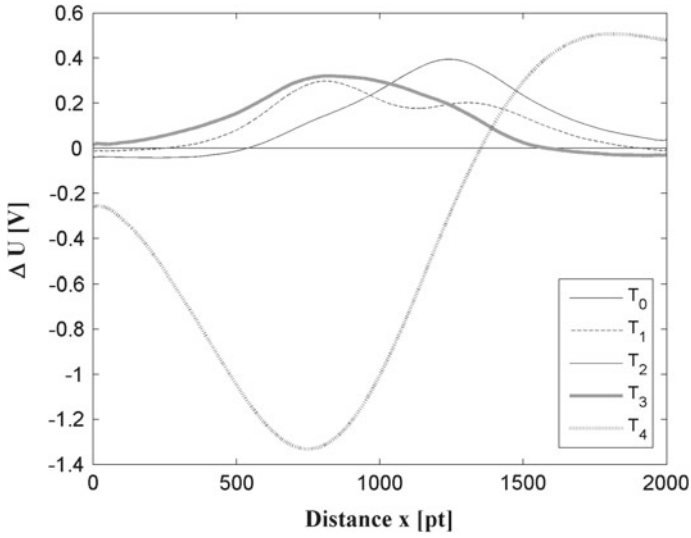


Fig. 8 Experimental results: maximum field values, AISI 304

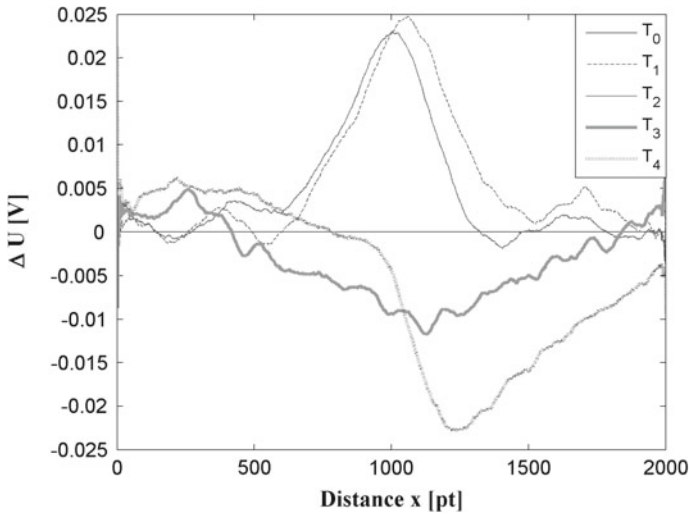


Fig. 9 Experimental results: maximum field values, AISI 316L

Obtained results for the AISI 316L are displayed in the Fig. 9. Their information content is comparable to the AISI 316Ti results. Further, individual waveforms have comparable information value. As it can be observed, it is quite difficult to find out differences between the waveforms, with respect to the signal values.

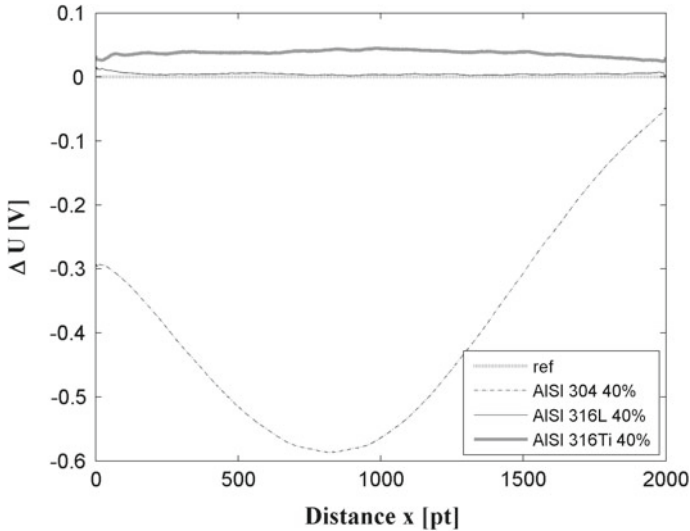


Fig. 10 Experimental results: lift-off $l = 20$ mm, maximum field values

Figure 10 display the waveforms, where the lift-off distance was set to $l = 20$ mm. This value was chosen to simulate the situation in real praxis: conductive biomaterial is located under the biological structure, in the distance of several millimeters. The only inspection is possible from this distance. The highest value of the plastic deformation ($T_4 = 40\%$) is compare in this graph. It can be observed that the highest signal strength has the material AISI 304 due to the presence of martensitic components. Further, it can be concluded that the also the AISI 316Ti material has different residual magnetic field values, in comparison with AISI 316L one. Based on the obtained results, it can be concluded: presence of the induced martensitic component can be evaluated using sensitive magnetic field sensors. The best results were achieved just for the material AISI 304, which possesses the highest measurable values of the magnetic field.

5 Conclusion

Impact of applied plastic deformation on residual magnetic field, in non-destructive evaluation of biomaterials, was investigated in the paper. The article presented magnetic field evaluation of the three different austenitic steel grades, used in medical praxis. Experimental measurements were carried out to compare the magnetic field of the individual specimens, before and after applied plastic deformation. All the measurements were performed under the same conditions. The presented results

showed that plastic deformation has strong impact on sensed responses. The residual magnetic field of such biomaterials can be evaluated using presented fluxgate sensor.

Further work of the authors will focus to detailed study of such biomaterials inspection, aimed to evaluation of fatigue cracking due to the phase transitions.

References

1. Boltz, E.S., Tiernan, T.C.: New electromagnetic sensors for detection of subsurface cracking and corrosion. *Rev. Prog. Quant. Non-Destruct. Eval.* **17**, 1033–1038 (1998)
2. Chen, Q., Thouas, G., A.: *Metallic implant biomaterials*. Material Science and Engineering R. Australia (2014)
3. Dogaru, T., Smith, S.T.: Edge crack detection using a giant magneto-resistance based eddy current sensor. *Non-Destruct. Test. Eval.* **16**, 31–53 (2000)
4. Grüger, H.: Array of miniaturized fluxgate sensors for non-destructive testing applications. www.sciencedirect.com. Accessed 03 Feb 2011
5. Kreutzbruck, M., Allweins, K., Heiden, C.: Fluxgate-magnetometer for the detection of deep lying defects. www.ndt.net. Accessed 28 Feb 2011
6. Lebrun, B., Jayet, Y., Baboux, J.C.: Pulsed eddy current application to the detection of deep cracks. *Mater. Eval.* **53**, 1296–1300 (1995)
7. Li, Y., Tian, G.Y., Simm, A.: Fast analytical modelling for pulsed eddy current evaluation. *NDT&E Int.* **41**, 477–483 (2008)
8. Palcek, P., Markovicova, L., Zatkalikova, V.: *Materials for Biomedical Engineering (Materialy pre biomedicinske inzinierstvo)*. University of Zilina, Zilina, EDIS (2015). (in Slovak)
9. Smetana, M., Capova, K., Palcek, P.: Austenitic steel magnetic properties evaluation after mechanical deformation, In: *Electromagnetic Nondestructive Evaluation (XVIII)*, pp. 35–42. IOS Press BV, Amsterdam (2015). ISBN 978-1-61499-508-1 (Studies in applied electromagnetics and mechanics, vol. 40. ISSN 1383-7281)
10. Smetana, M., Capova, K., Palcek, P., Oravcova, M.: Identification of biomaterial fatigue cracks by ECT method. In: *Proceedings of the 20th International Workshop on Electromagnetic Non-destructive Evaluation*. Sendai, Japan (2015)
11. Ward, W.W., Moulder, J.C.: Lowfrequency, pulsed eddy currents for deep penetration. *Rev. Prog. Quant. Non-Destruct. Eval.* **17**, 291–298 (1998)

Author Index

A

Antoniak, Kinga, [15](#)
Antonowicz, Magdalena, [217](#)
Aumayr, Georg, [3](#)

B

Badura, Paweł, [63](#)
Bednorz, Adam, [15](#)
Benova, Mariana, [245](#)
Blahovcová, Eva, [207](#)
Borys, Damian, [227](#)
Bugdol, Marcin D., [41](#)
Bugdol, Monika N., [63](#)
Burkacki, Michał, [299](#)

C

Capova, Klara, [399](#)
Chuchnowska, Iwona, [99](#)
Chudacik, Vladimír, [399](#)
Conradt, Jörg, [87](#)
Czajkowska, Joanna, [235](#)

D

Danch-Wierzchowska, Marta, [227](#)
Derejczyk, Jarosław, [15](#)

F

Foszner, Paweł, [161](#)
Fujarewicz, Krzysztof, [123](#), [173](#)
Furmaniak, Joanna, [49](#)

G

Gertych, Arkadiusz, [193](#)
Glonek, Grzegorz, [307](#)
Grygiel-Pradelok, Magdalena, [385](#)
Grzegorzek, Marcin, [27](#), [41](#)
Guzik-Kopyto, Agata, [99](#)
Gzik, Marek, [289](#), [299](#)
Gzik-Zroska, Bożena, [289](#), [299](#)

H

Halašová, Erika, [207](#)
Hanusiak, Aneta, [15](#)
Hippe, Zdzisław S., [363](#)
Hoffmann, Raoul, [87](#)
Hyla, Agnieszka, [375](#)

I

Ing, Nathan, [193](#)

J

Jakubczak, Michał, [123](#)
Jochymczyk-Woźniak, Katarzyna, [75](#)
Joszko, Kamil, [289](#), [299](#)
Jurkojć, Jacek, [75](#)
Juszczuk, Jan, [235](#)

K

Kaczmarek, Marcin, [255](#)
Kaczorowski, Roman, [49](#)
Kajzer, Anita, [217](#), [385](#)
Kajzer, Wojciech, [217](#), [385](#)

Kardynska, Malgorzata, 275
Kasprzycki, Maciej, 49
Kawa, Jacek, 15
Knudsen, Beatrice, 193
Koktysz, Robert, 183
Kopyta, Ilona, 75
Kowalski, Przemysław, 267
Kozłowski, Wojciech, 183
Kręcichwost, Michał, 63
Kubiak, Zygmunt, 49
Kupkova, Kristyna, 133
Kurtyka, Przemysław, 255

L

Lach, Ewa, 337
Lach, Joanna, 15
Lauterbach, Christl, 87
Luchowski, Leszek, 267
Łukasik, Ewa, 49

M

Ma, Zhaoxuan, 193
Mandera, Marek, 75
Markiewicz, Tomasz, 183
Meyer, Robert, 41
Michnik, Robert, 75, 99
Miodońska, Zuzanna, 63
Mištuna, Dušan, 207
Mitas, Andrzej W., 41

N

Nawrat-Szoltysik, Agnieszka, 63
Nowakowska, Katarzyna, 75

O

Oravcova, Monika, 399

P

Pawlyta, Magdalena, 321
Pezowicz, Celina, 349
Pieter, Justyna, 173
Piętka, Ewa, 235
Polanski, Andrzej, 161
Provaznik, Ivo, 133
Psenakova, Zuzana, 245
Pycinski, Bartłomiej, 235

R

Rudzki, Marcin, 63

S

Salman, Sadri, 193
Sciuto, Grazia Lo, 245
Sedlar, Karel, 133
Seiffert, Piotr, 15
Shirahama, Kimiaki, 27
Skabek, Krzysztof, 267
Škovierová, Henrieta, 207
Skurowski, Przemysław, 321
Smetana, Milan, 399
Smieja, Jarosław, 275
Smondrk, Maros, 245
Sobczyk, Karolina, 173
Steinhage, Axel, 87
Stepien, Paula, 63
Strnádel, Ján, 207
Student, Sebastian, 173
Suchoń, Sławomir, 299
Swiderska-Chadaj, Zaneta, 183
Swierniak, Andrzej, 227
Szmyszal, Jan, 15

T

Tamulewicz, Anna, 147
Tarnawski, Michał, 267
Techmer, Axel, 87
Tkacz, Ewaryst, 147
Tomaka, Agnieszka Anna, 267

U

Utracki, Jarosław, 109

W

Walke, Witold, 255, 375
Walts, Ann, 193
Wilhelm, Christoph, 41
Wodarski, Piotr, 99
Wojciechowski, Adam, 307
Wolański, Wojciech, 289, 299

Z

Zarychta, Piotr, 63
Zgrzeba, Bartosz, 49
Ziębowicz, Anna, 385
Ziębowicz, Bogusław, 385
Żuk, Magdalena, 349