

Ignacio Rojas
Francisco Ortuño (Eds.)

LNBI 10814

Bioinformatics and Biomedical Engineering

6th International Work-Conference, IWBBIO 2018
Granada, Spain, April 25–27, 2018
Proceedings, Part II

2 Part II

 Springer

Subseries of Lecture Notes in Computer Science

LNBI Series Editors

Sorin Istrail

Brown University, Providence, RI, USA

Pavel Pevzner

University of California, San Diego, CA, USA

Michael Waterman

University of Southern California, Los Angeles, CA, USA

LNBI Editorial Board

Søren Brunak

Technical University of Denmark, Kongens Lyngby, Denmark

Mikhail S. Gelfand

IITP, Research and Training Center on Bioinformatics, Moscow, Russia

Thomas Lengauer

Max Planck Institute for Informatics, Saarbrücken, Germany

Satoru Miyano

University of Tokyo, Tokyo, Japan

Eugene Myers

Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Marie-France Sagot

Université Lyon 1, Villeurbanne, France

David Sankoff

University of Ottawa, Ottawa, Canada

Ron Shamir

Tel Aviv University, Ramat Aviv, Tel Aviv, Israel

Terry Speed

Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia

Martin Vingron

Max Planck Institute for Molecular Genetics, Berlin, Germany

W. Eric Wong

University of Texas at Dallas, Richardson, TX, USA

More information about this series at <http://www.springer.com/series/5381>

Ignacio Rojas · Francisco Ortuño (Eds.)

Bioinformatics and Biomedical Engineering

6th International Work-Conference, IWBBIO 2018
Granada, Spain, April 25–27, 2018
Proceedings, Part II

Editors
Ignacio Rojas
University of Granada
Granada
Spain

Francisco Ortuño
University of Granada
Granada
Spain

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Bioinformatics
ISBN 978-3-319-78758-9 ISBN 978-3-319-78759-6 (eBook)
<http://doi.org/10.1007/978-3-319-78759-6>

Library of Congress Control Number: 2018937390

LNCS Sublibrary: SL8 – Bioinformatics

© Springer International Publishing AG, part of Springer Nature 2018, corrected publication 2019
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. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

We are proud to present the set of final accepted full papers for the third edition of the IWBBIO conference “International Work-Conference on Bioinformatics and Biomedical Engineering” held in Granada (Spain) during April 25–27, 2018.

The IWBBIO 2018 (International Work-Conference on Bioinformatics and Biomedical Engineering) conference sought to provide a discussion forum for scientists, engineers, educators, and students about the latest ideas and realizations in the foundations, theory, models, and applications for interdisciplinary and multidisciplinary research encompassing disciplines of computer science, mathematics, statistics, biology, bioinformatics, and biomedicine.

The aims of IWBBIO 2018 were to create a friendly environment that could lead to the establishment or strengthening of scientific collaborations and exchanges among attendees, and therefore, IWBBIO 2018 solicited high-quality original research papers (including significant work-in-progress) on any aspect of bioinformatics, biomedicine, and biomedical engineering.

We especially encouraged contributions dealing with new computational techniques and methods in machine learning; data mining; text analysis; pattern recognition; data integration; genomics and evolution; next-generation sequencing data; protein and RNA structure; protein function and proteomics; medical informatics and translational bioinformatics; computational systems biology; modelling and simulation and their application in life science domain, biomedicine, and biomedical engineering. The list of topics in the successive call for papers also evolved, resulting in the following list for the present edition:

1. Computational proteomics. Analysis of protein–protein interactions. Protein structure modelling. Analysis of protein functionality. Quantitative proteomics and PTMs. Clinical proteomics. Protein annotation. Data mining in proteomics.
2. Next-generation sequencing and sequence analysis. De novo sequencing, re-sequencing and assembly. Expression estimation. Alternative splicing discovery. Pathway Analysis. Chip-seq and RNA-Seq analysis. Metagenomics. SNPs prediction.
3. High performance in bioinformatics. Parallelization for biomedical analysis. Biomedical and biological databases. Data mining and biological text processing. Large-scale biomedical data integration. Biological and medical ontologies. Novel architecture and technologies (GPU, P2P, Grid) for Bioinformatics.
4. Biomedicine. Biomedical computing. Personalized medicine. Nanomedicine. Medical education. Collaborative medicine. Biomedical signal analysis. Biomedicine in industry and society. Electrotherapy and radiotherapy.
5. Biomedical engineering. Computer-assisted surgery. Therapeutic engineering. Interactive 3D modelling. Clinical engineering. Telemedicine. Biosensors and data acquisition. Intelligent instrumentation. Patient Monitoring. Biomedical robotics. Bio-nanotechnology. Genetic engineering.

6. Computational systems for modelling biological processes. Inference of biological networks. Machine learning in bioinformatics. Classification for biomedical data. Microarray data analysis. Simulation and visualization of biological systems. Molecular evolution and phylogenetic modelling.
7. Health care and diseases. Computational support for clinical decisions. Image visualization and signal analysis. Disease control and diagnosis. Genome-phenome analysis. Biomarker identification. Drug design. Computational immunology.
8. E-health. E-health technology and devices. E-health information processing. Telemedicine/E-health application and services. Medical image processing. Video techniques for medical images. Integration of classical medicine and e-health.

After a careful peer review and evaluation process (each submission was reviewed by at least two, and on average 3.1, Program Committee members or additional reviewer), 88 papers were accepted for oral, poster, or virtual presentation, according to the recommendations of the reviewers and the authors' preferences, and to be included in the LNBI proceedings.

During IWBBIO 2018, several Special Sessions were held. Special Sessions are a very useful tool for complementing the regular program with new and emerging topics of particular interest for the participating community. Special Sessions that emphasize multidisciplinary and transversal aspects, as well as cutting-edge topics, were especially encouraged and welcomed, and in this edition of IWBBIO 2018 were the following:

– **SS1. Generation, Management, and Biological Insights from Big Data.**

As the sequencing technologies develop, reducing the costs and increasing the accuracy, research in biological sciences is transformed from hypothesis-driven to data-driven approaches. Big data encompasses a generation of data ranging from DNA sequence data for thousands of individuals to single-cell data for thousands of cell types from an individual. This has moved the bottle-neck of the data generation down-stream to use these data to gain new knowledge, finally with an aim to improve the quality of human life. The important down-stream challenges with big data include development of strategies for efficient storage of big data making them findable, accessible, interoperable, and reusable (FAIR), to make them usable for research. The next step is the development of new methods including software and Web tools to make sense of big data. The final important step is to demonstrate that big data can indeed lead to new knowledge. This session will cover the research topics in all three aspects of big data described here.

Organizer: Dr. Anagha Joshi, Group leader in the Division of Developmental Biology at the Roslin Institute, University of Edinburgh, UK.

Website:

<https://www.ed.ac.uk/roslin/about/contact-us/staff/anagha-joshi>

– **SS2. Challenges in Smart and Wearable Sensor Design for Mobile Health.**

The analysis of data streams captured with biomedical sensors can be performed as an embedded procedure within the sensor or sensor network or at a later stage in a

receiving system. Currently, several systems reduce the number of signals monitored via sensors (e.g., when using wearable devices) in order to save energy. In this case, the pre-processing task is embedded into the sensor or close to it. As a result, fewer data are transferred but pattern matching becomes more complex since cross-reference data are missing and computing power is limited. This session should present new and emerging approaches.

Organizers: Prof. Natividad Martínez Madrid, Head of the Internet of Things Laboratory and Director of the AAL-Living Lab at Reutlingen University, Germany.

Prof. Juan Antonio Ortega, Director of the Centre of Computer Scientific in Andalusia (Spain) www.cica.es and head of the research group IDINFOR (TIC223), University of Seville, ETS Ingeniería Informática, Spain.

Prof. Ralf Seepold, Head of the Ubiquitous Computing Lab at HTWG Konstanz, Department of Computer Science, Germany.

Websites:

<http://iotlab.reutlingen-university.de>

<http://madeirasic.us.es/idinfor/>

<http://uc-lab.in.htwg-konstanz.de>

– SS3. Challenges and Advances in Measurement and Self-Parametrization of Complex Biological Systems.

Our understanding of biological systems requires progress in the measurement techniques, methods, and principles of acquisition. The development of IT and physical resolution offers novel advanced probes, devices, or interpretation as well as more questions and possibilities. Automation of processing and analysis is increasing thanks to artificial intelligence and machine deep learning. The proper bioinformatic parametrization for the analysis of complex systems continues toward automatic or self-setting of the acquired biophysical attributes.

This special section provided a discussion on novel techniques and measurement devices, emerging challenges for complex systems, open solutions, and future visions. The broad examples from self-parametric results supported the discussion with practical applications.

Organizer: Dipl.-Ing. Jan Urban, PhD, Head of Laboratory of Signal and Image Processing, University of South Bohemia in České Budejovice, Faculty of Fisheries and Protection of Waters, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Institute of Complex Systems, Czech Republic.

Website:

www.frov.jcu.cz/en/institute-complex-systems/lab-signal-image-processing

– SS4. High-Throughput Bioinformatic Tools for Medical Genomics.

Genomics is concerned with the sequencing and analysis of an organism's genome taking advantage of the current, cost-effective, high-throughput sequencing

technologies. Their continuous improvement is creating a new demand for enhanced high-throughput bioinformatics tools. In this context, the generation, integration, and interpretation of genetic and genomic data are driving a new era of health care and patient management. Medical genomics (or genomic medicine) is this emerging discipline that involves the use of genomic information about a patient as part of the clinical care with diagnostic or therapeutic purposes to improve the health outcomes. Moreover, it can be considered a subset of precision medicine that is having an impact in the fields of oncology, pharmacology, rare and undiagnosed diseases, and infectious diseases. The aim of this special session is to bring together researchers in medicine, genomics, and bioinformatics to translate medical genomics research into new diagnostic, therapeutic, and preventive medical approaches. Therefore, we invited authors to submit original research, new tools or pipelines, or their update, and review articles on relevant topics, such as (but not limited to):

- Tools for data pre-processing (quality control and filtering)
- Tools for sequence mapping
- Tools for the comparison of two read libraries without an external reference
- Tools for genomic variants (such as variant calling or variant annotation)
- Tools for functional annotation: identification of domains, orthologues, genetic markers, controlled vocabulary (GO, KEGG, InterPro)
- Tools for biological enrichment in non-model organisms
- Tools for gene expression studies
- Tools for Chip-Seq data
- Tools for “big data” analyses
- Tools for integration in workflows

Organizers: Prof. M. Gonzalo Claros, Department of Molecular Biology and Biochemistry, University of Málaga, Spain.

Dr. Javier Pérez Florido, Bioinformatics Research Area, Fundación Progreso y Salud, Seville, Spain.

– **SS5. Drug Delivery System Design Aided by Mathematical Modelling and Experiments.**

A drug delivery system is designed to release controlled amount of drugs to a specific target area. To devise optimization strategies for targeted drug delivery the combined action of various processes needs to be well understood. Mathematical modelling offers a valuable tool when evaluating potential drug-carrying materials coupled with rate-controlling coatings. When applied to experimental data, simulations can yield valuable insight and guide further research with the aim of identifying and evaluating key drug release mechanisms. Although diffusion is often a primary drug release process, other effects such as binding and dissolution as well as effects occurring at material interfaces are no less important in describing various rate-controlling release mechanisms.

Considered systems include: intraocular- and soft contact lenses, orthopedic implants, arterial stents, and transdermal patches.

Organizers: PhD candidate Kristinn Gudnason, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Iceland.

Prof. Fjola Jonsdottir, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Iceland.

Prof. Emeritus Sven Sigurdsson, Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Iceland.

Prof. Mar Masson, Faculty of Pharmaceutical Science, University of Iceland, Iceland.

– **SS6. Molecular Studies on Inorganic Nanomaterials for Therapeutical and Diagnosis Applications.**

Nanostructured material science with natural origins is becoming a hot topic in nanomedicine for addressing toxicity and high cost limitations. The absorption of pharmaceutical drugs in natural inorganic nanostructured solids is very useful for controlled delivery of bioactive compounds. Molecular modelling and analytical spectroscopic techniques are well-established research fields for the characterization of these materials. This approach is becoming of great interest in the studies of these nanocomposites and the interactions of organics on the surfaces of the inorganic solids in health applications. The aim of this session is to gather professionals from a wide scope of scientific disciplines to better understand molecular aspects of nanocomposite components behavior and drug design. This interdisciplinary session included contributions from computational chemistry (empirical potentials, quantum, coarse-grained, etc.), NMR, infrared, and Raman spectroscopies, as well as X-ray-diffraction/neutron/synchrotron techniques.

This special session has a multidisciplinary nature and is not easy to be included in one of the congress topics owing to its transversal aim connected with several topics of the congress: computational proteomics (protein structure modelling), biomedicine (biomedical computing, nanomedicine), biomedical engineering (bio-nanotechnology), computational systems for modelling biological processes (simulation and visualization of biological systems), health care and diseases (drug design and computational immunology).

The aim of this special session is to show the potential application of computational modelling methods in nanomedicine for experimental researchers and, at the same time, for theoreticians diagnosing possible complementary tools for experiments, generating useful discussions between experimentalists and theoreticians to promote future scientific collaborations.

Organizers: Dr. C. Ignacio Sainz-Díaz, Instituto Andaluz de Ciencias de la Tierra, CSIC/UGR, Granada, Spain.

Dr. Carola Aguzzi, Departamento de Tecnología Farmacéutica, Universidad de Granada, Granada, Spain.

– **SS7. Little-Big Data. Reducing the Complexity and Facing Uncertainty of Highly Underdetermined Phenotype Prediction Problems.**

Phenotype prediction problems have a very underdetermined character since the number of samples is always much lower in size than the number of genes/genetic probes/SNPs/etc., that are monitored to explain a given phenotype. This generates decision problems that have a huge uncertainty space. This includes a great variety of problems with great impact in translational medicine, such as the analysis of mechanism of action of genes in disease progression, the investigation of new therapeutic targets, the analysis of secondary effects, treatment optimization, as well as analysis of the effect of mutations in the transcriptome and in proteomics, etc. The objective of the session was to present novel computational approaches to reduce the complexity of high-dimensional genetic data while keeping the main information content. Applications in cancer and genomics as well as rare and neurodegenerative diseases were welcome. In particular, the design of new methods to perform the robust analysis of pathways involved in disease development were one of the main topics addressed in this session.

Organizer: Prof. Juan Luis Fernández-Martínez, Mathematics Department, Applied Mathematics Section, Director of the Group of Inverse Problems, Optimization and Machine Learning, University of Oviedo, Spain.

– **SS8. Interpretable Models in Biomedicine and Bioinformatics.**

In a very short period of time, many areas of science have made a sharp transition toward data-driven methods. This new situation is clear in the life sciences and, as particular cases, in biomedicine, bioinformatics, and health care.

You could see this as a perfect scenario for the use of data analytics, from multivariate statistics to machine learning (ML) and computational intelligence (CI), but this scenario also poses some serious challenges. One of them takes the form of (lack of) interpretability/comprehensibility/explainability of the models obtained through data analysis. This could be a bottleneck especially for complex nonlinear models, often affected by what has come to be known as the “black box syndrome.” In some areas such as medicine and health care, not addressing such challenges might seriously limit the chances of adoption, in real practice, of computer-based medical decision support systems (MDSS).

Interpretability and explainability have become hot research issues, and there are different reasons for this: One of them is the soaring success of deep learning artificial neural networks in recent years. These models risk not being adopted in areas where human decision is key and that decision must be explained as they are extreme “black box” cases. Another reason is the implementation of the European Union directive for General Data Protection Regulation (GDPR). Enforced in April 2018, it mandates a right to explanation of all decisions made by automated or artificially intelligent algorithmic systems. Needless to say, this directly involves data analytics and it is likely to have an impact on health care, medical

decision-making, and even in bioinformatics through the use of genomics in personalized medicine.

In this session, we called for papers that broach the topics of interpretability/comprehensibility/explainability of data models (with a non-reductive focus on ML and CI) in biomedicine, bioinformatics, and health care, from different viewpoints, including:

- Enhancement of the interpretability of existing data analysis techniques in problems related to biomedicine, bioinformatics, and health care
- New methods of model interpretation/explanation in problems related to biomedicine, bioinformatics, and health care
- Case studies biomedicine, bioinformatics, and health care in which interpretability/comprehensibility/explainability is a key aspect of the investigation
- Methods to enhance interpretability in safety-critical areas (such as, for instance, critical care)
- Issues of ethics and social responsibility (including governance, privacy, anonymization) in biomedicine, bioinformatics, and health care

Organizers: Prof. Alfredo Vellido, Intelligent Data Science and Artificial Intelligence (IDEAI) Research Center, Universitat Politècnica de Catalunya, Barcelona, Spain.

Prof. Sandra Ortega-Martorell, Department of Applied Mathematics, Liverpool John Moores University, Liverpool, UK.

Prof. Alessandra Tosi, Mind Foundry Ltd., Oxford, UK.

Prof. Iván Olier Caparros, MMU Machine Learning Research Lab, Manchester Metropolitan University, Manchester, UK.

– SS9. Medical Planning: Management System for Liquid Radioactive Waste in Hospital Design.

In tertiary hospitals where nuclear medicine services have been introduced, the radioactive materials used in diagnosis and/or treatment need to be handled. The hospital design and medical planning should consider these materials and their policy for treatment. Nuclear waste has been divided into solid and liquid based on the materials used and on their half-life times, which start from a few minutes to years.

In our study, the most common radioactive liquid materials (waste) were treated by smart systems that detect the material and based on its HLT (activities) will be distributed in shielded storage tanks to the sewage treatment plant (STP) of the hospital after keeping them for the required times. The location and capacity of these tanks together with their monitoring and control system should be considered in the design stage that determines the treatment processes.

Motivation and objectives for the session: The nuclear medicine department should be considered in the design stage and its space program. The location and capacity of storage tanks and their drainage lines should be considered in hospital.

Organizer: Dr. Khaled El-Sayed, Assistant Professor of Biomedical Engineering, Department of Electrical and Medical Engineering, Benha University, Egypt.

– **SS10. Bioinformatics Tools to Integrate Omics Dataset and Address Biological Question.**

Methodological advances in ‘omics’ technologies allow for the high-throughput detection and monitoring of the abundance of several biological molecules. Several ‘omics’ platforms are being used in molecular biology and in clinical practice for the understanding of molecular mechanisms underlying specific disease as well as for identifying trustworthy diagnostic/prognostic markers.

Omics strategies include: genomics, which aims to characterize and quantify a set of genes within a single cell of an organism; transcriptomics, which analyzes the levels of mRNA transcripts; proteomics, which includes the identification of proteins and the monitoring of their abundance; metabolomics, which measures the abundance of small cellular metabolites; interactomics, and many others. Although more informative, no single ‘omics’ analysis may fully unveil the complexities of a specific biological question, therefore, to achieve a more comprehensive “picture” of biological processes, experimental data made on different layers have to be integrated and analyzed. This special session aimed to provide a description of bioinformatics strategies aimed at integrating omics datasets to address biological questions.

Organizer: Dr. Domenica Scumaci, PhD Laboratory of Proteomics, Department of Experimental and Clinical Medicine, Magna Graecia University of Catanzaro, Italy.

– **SS11. Understanding the Mechanisms of Variant Effects on Human Disease Phenotype.**

Modern sequencing technologies have enabled whole-genome sequencing and detailed quantification of germline and somatic variations, many of which are related to human disease. However, these data are necessary but not sufficient for understanding the cause and effect of these variations as well as their mechanisms of involvement in human diseases phenotypes. Disease-related variants can have an impact on DNA, RNA, and protein functions and may lead to impaired replication, transcription, signal transduction, and epigenetic regulation. This session exploited computational approaches related to inferring the effects of human mutations on proteins, biomolecular interactions, and cellular pathways with the goal of elucidating mechanistic aspects of disease causative variants.

Organizer: Anna Panchenko, PhD Head, Computational Biophysics Group Computational Biology Branch, National Center for Biotechnology Information, National Institutes of Health, Bethesda, USA.

Website:

<https://www.ncbi.nlm.nih.gov/CBBResearch/Panchenko/>

In this edition of IWBBIO, we were honored to have the following invited speakers:

1. Prof. Joaquin Dopazo, Fundacion Progreso y Salud, Clinical Bioinformatics Research Area, Seville, Spain
2. Prof. Luis Rueda, School of Computer Science, Pattern Recognition and Bioinformatics Lab, Windsor Cancer Research Group, University of Windsor
3. Dr. Anagha Joshi, Bioinformatics Group Leader, Developmental Biology Division, The Roslin Institute, University of Edinburgh, UK
4. Prof. FangXiang Wu, SMIEEE Professor, Division of Biomedical Engineering, Department of Mechanical Engineering, College of Engineering, University of Saskatchewan, Canada
5. Prof. Jiayin Wang, Xi an Jiaotong University, China

It is important to note, that for the sake of consistency and readability of the book, the presented papers are classified under 17 chapters. The organization of the papers is in two volumes arranged following the topics list included in the call for papers. The first volume (LNBI 10813), entitled *Advances in Computational Intelligence. Part I* is divided into 11 main parts and includes the contributions on:

1. Bioinformatics for health care and diseases
2. Bioinformatics tools to integrate omics datasets and address biological questions
3. Challenges and advances in measurement and self-parametrization of complex biological systems
4. Computational genomics
5. Computational proteomics
6. Computational systems for modelling biological processes
7. Drug delivery system design aided by mathematical modelling and experiments
8. Generation, management, and biological insights from big data
9. High-throughput bioinformatic tools for medical genomics
10. Next-generation sequencing and sequence analysis
11. Interpretable models in biomedicine and bioinformatics

In the second volume (LNBI 10814), entitled *Advances in Computational Intelligence. Part II* is divided into six main parts and includes the contributions on:

1. Little-big data. Reducing the complexity and facing uncertainty of highly underdetermined phenotype prediction problems
2. Biomedical engineering
3. Biomedical image analysis
4. Biomedical signal analysis
5. Challenges in smart and wearable sensor design for mobile health
6. Health care and diseases

This sixth edition of IWBBIO was organized by the Universidad de Granada together with the Spanish Chapter of the IEEE Computational Intelligence Society. We

wish to thank to our main sponsor and the Faculty of Science, Department of Computer Architecture and Computer Technology, and CITIC-UGR from the University of Granada for their support and grants. We wish also to thank to the Editors-in-Chief of different international journals for their interest in editing special issues of the best papers of IWBBIO.

We would also like to express our gratitude to the members of the different committees for their support, collaboration, and good work. We especially thank the local Organizing Committee, Program Committee, the reviewers, and special session organizers. We also want to express our gratitude for the EasyChair platform. Finally, we want to thank Springer, and especially Alfred Hoffman and Anna Kramer for their continuous support and cooperation.

April 2018

Ignacio Rojas
Francisco Ortuño

Organization

Steering Committee

Miguel A. Andrade	University of Mainz, Germany
Hesham H. Ali	University of Nebraska, USA
Oresti Baños	University of Twente, The Netherlands
Alfredo Benso	Politecnico di Torino, Italy
Giorgio Buttazzo	Superior School Sant'Anna, Italy
Gabriel Caffarena	University San Pablo CEU, Spain
Mario Cannataro	University Magna Graecia of Catanzaro, Italy
Jose María Carazo	Spanish National Center for Biotechnology (CNB), Spain
Jose M. Cecilia	Universidad Católica San Antonio de Murcia (UCAM), Spain
M. Gonzalo Claros	University of Malaga, Spain
Joaquin Dopazo	Research Center Principe Felipe (CIPF), Spain
Werner Dubitzky	University of Ulster, UK
Afshin Fassihi	Universidad Católica San Antonio de Murcia (UCAM), Spain
Jean-Fred Fontaine	University of Mainz, Germany
Humberto Gonzalez	University of Basque Country (UPV/EHU), Spain
Concettina Guerra	College of Computing, Georgia Tech, USA
Roderic Guigo	Center for Genomic Regulation, Pompeu Fabra University, Spain
Andy Jenkinson	Karolinska Institute, Sweden
Craig E. Kapfer	Reutlingen University, Germany
Narsis Aftab Kiani	European Bioinformatics Institute (EBI), UK
Natividad Martinez	Reutlingen University, Germany
Marco Masseroli	Polytechnic University of Milan, Italy
Federico Moran	Complutense University of Madrid, Spain
Cristian R. Munteanu	University of Coruña, Spain
Jorge A. Naranjo	New York University (NYU), Abu Dhabi
Michael Ng	Hong Kong Baptist University, SAR China
Jose L. Oliver	University of Granada, Spain
Juan Antonio Ortega	University of Seville, Spain
Julio Ortega	University of Granada, Spain
Alejandro Pazos	University of Coruña, Spain
Javier Perez Florido	Genomics and Bioinformatics Platform of Andalusia, Spain
Violeta I. Pérez Nueno	Inria Nancy Grand Est (LORIA), France
Horacio Pérez-Sánchez	Universidad Católica San Antonio de Murcia (UCAM), Spain

Alberto Policriti	Università di Udine, Italy
Omer F. Rana	Cardiff University, UK
M. Francesca Romano	Superior School Sant'Anna, Italy
Yvan Saey	VIB - Ghent University
Vicky Schneider	The Genome Analysis Centre (TGAC), UK
Ralf Seepold	HTWG Konstanz, Germany
Mohammad Soruri	University of Birjand, Iran
Yoshiyuki Suzuki	Tokyo Metropolitan Institute of Medical Science, Japan
Oswaldo Trelles	University of Malaga, Spain
Shusaku Tsumoto	Shimane University, Japan
Renato Umeton	CytoSolve Inc., USA
Jan Urban	University of South Bohemia, Czech Republic
Alfredo Vellido	Polytechnic University of Catalonia, Spain
Wolfgang Wurst	GSF National Research Center of Environment and Health, Germany

Program Committee and Additional Reviewers

Jesus S. Aguilar	Christina Boucher
Carlos Alberola	Hacene Boukari
Hisham Al-Mubaid	Daniel Brown
Rui Carlos Alves	Fiona Browne
Yuan An	Dongbo Bu
Georgios Anagnostopoulos	Jeremy Buhler
Eduardo Andrés León	Keith C. C.
Antonia Aránega	Carlos Cano
Saúl Ares	Angel Cantu
Masanori Arita	Rita Casadio
Ruben Armañanzas	Daniel Castillo
Joel P. Arrais	Oswaldo Castellanos
Patrizio Arrigo	Ting-Fung Chan
O. Bamidele Awojoyogbe	Nagasuma Chandra
Hazem Bahig	Kun-Mao Chao
Pedro Ballester	Bolin Chen
Graham Balls	Brian Chen
Ugo Bastolla	Chuming Chen
Sidahmed Benabderrahmane	Jie Chen
Steffanny A. Bennett	Yuehui Chen
Alfredo Benso	Jianlin Cheng
Mahua Bhattacharya	Shuai Cheng
Concha Bielza	I-Jen Chiang
Armando Blanco	Jung-Hsien Chiang
Ignacio Blanquer	Young-Rae Cho
Olivier Bodenreider	Justin Choi
Paola Bonizzoni	Petr Cisar

Darrell Conklin
Clare Coveney
Aedin Culhane
Miguel Damas
Bhaskar DasGupta
Ricardo De Matos
Guillermo de la Calle
Javier De Las Rivas
Fei Deng
Marie-Dominique Devignes
Sergio Diaz-Del-Pino
Ramón Diaz-Uriarte
Julie Dickerson
Ye Duan
Beatrice Duval
Khaled El-Sayed
Mamdoh Elsheshengy
Christian Exposito
Weixing Feng
Jose Jesús Fernandez
Gionata Fragomeni
Xiaoyong Fu
Alexandre G. de Brevern
Eduardo Gade Gusmao
Juan Manuel Galvez
Pugalenthi Ganesan
Jean Gao
Qingsong Gao
Rodolfo Garcia
Lina Gaudio
Mark Gerstein
Razvan Ghinea
Daniel Gonzalez Peña
Dianjing Guo
Jun-tao Guo
Maozu Guo
Christophe Guyeux
Michael Hackenberg
Michiaki Hamada
Xiyi Hang
Jin-Kao Hao
Nurit Haspel
Morihiro Hayashida
Jieyue He
Luis Javier Herrera
Pietro Hiram
Lynette Hirschman
Ralf Hofestadt
Vasant Honavar
Jun Hu
Xiaohua Hu
Jun Huan
Chun-Hsi Huang
Heng Huang
Jimmy Huang
Jingshan Huang
Jianzheng Huang
Seiya Imoto
Jiri Jablonsky
Guomin Ji
Yanqing Ji
Xingpeng Jiang
Chandra Kambhamettu
Mingon Kang
Dong-Chul Kim
Dongsup Kim
Hyunsoo Kim
Sun Kim
Kengo Kinoshita
Ekaterina Kldiashvili
Jun Kong
Tomas Koutny
Natalio Krasnogor
Abhay Krishan
Marija Krstic-Demonacos
Stephen Kwok-Wing
Istvan Ladunga
T. W. Lam
Jorge Langa
Dominique Lavenier
Jose Luis Lavin
Doheon Lee
Xiujuan Lei
André Leier
Kwong-Sak Leung
Chen Li
Dingcheng Li
Jing Li
Jinyan Li
Min Li
Xiaoli Li
Yanpeng Li

Li Liao
Hongfei Lin
Hongfang Liu
Jinze Liu
Xiaowen Liu
Xiong Liu
Zhenqiu Liu
Zhi-Ping Liu
Rémi Longuespée
Miguel Angel Lopez Gordo
Ernesto Lowy
Jose Luis
Suryani Lukman
Feng Luo
Qin Ma
Malika Mahoui
Tatiana Marquez-Lago
Keith Marsolo
Francisco Martinez Alvarez
Tatiana Maximova
Roderik Melnik
Pall Melsted
Jordi Mestres
Hussain Michelle
Marianna Milano
Ananda Mondal
Antonio Morreale
Walter N. Moss
Maurice Mulvenna
Enrique Muro
Radhakrishnan Nagarajan
Vijayaraj Nagarajan
Kenta Nakai
Isabel A. Nepomuceno
Mohammad Nezami
Anja Nohe
Michael Ochs
Baldomero Oliva
Jose Luis Oliveira
Motonori Ota
David P.
Tun-Wen Pai
Paolo Paradisi
Hyun-Seok Park
Kunsoo Park
Taesung Park
David Pelta
Alexandre Perera
María Del Mar Pérez Gómez
Esteban Perez-Wohlfeil
Vinhthuy Phan
Antonio Pinti
Héctor Pomares
Mihail Popescu
Benjarath Pupacdi
Sanguthevar Rajasekaran
Shoba Ranganathan
Patrick Riley
Jairo Rocha
Fernando Rojas
Jianhua Ruan
Gregorio Rubio
Antonio Rueda
Irena Rusu
Renata Rychtarikova
Vincent Shin-Mu Tseng
Mohammadmehdi Saberioon
Kunihiko Sadakane
Michael Sadovsky
Belen San Roman
Maria Jose Saez
Hiroto Saigo
José Salavert
Carla Sancho Mestre
Emmanuel Sapin
Kengo Sato
Jean-Marc Schwartz
Russell Schwartz
Jose Antonio Seoane
Xuequn Shang
Piramanayagam Shanmughavel
Xiaoman Shawn
Xinghua Shi
Tetsuo Shibuya
Tiratha Raj Singh
Dong-Guk Shin
Amandeep Sidhu
Istvan Simon
Richard Sinnott
Jiangning Song
Zhengchang Su
Joakim Sundnes

Wing-Kin Sung
Prashanth Suravajhala
Martin Swain
Sing-Hoi Sze
Mehmet Tan
Xing Tan
Li Teng
Dang Thanh
Tianhai Tian
Pedro Tomas
Carlos Toro
Carolina Torres
Paolo Trunfio
Esko Ukkonen
Olga Valenzuela
Lucia Vaira
Paola Velardi
Julio Vera
Konstantinos Votis
Slobodan Vucetic
Ying-Wooi Wan
Chong Wang
Haiying Wang
Jason Wang
Jialin Wang
Jian Wang
Jianxin Wang
Jiayin Wang
Junbai Wang
Junwen Wang
Lipo Wang
Lusheng Wang
Yadong Wang
Yong Wang

Ka-Chun Wong
Ling-Yun Wu
Xintao Wu
Zhonghang Xia
Fang Xiang
Lei Xu
Zhong Xue
Patrick Xuechun
Hui Yang
Zhihao Yang
Jingkai Yu
Hong Yue
Erliang Zeng
Xue-Qiang Zeng
Aidong Zhang
Chi Zhang
Jiao Zhang
Jin Zhang
Jingfen Zhang
Kaizhong Zhang
Le Zhang (Adam)
Shao-Wu Zhang
Xingan Zhang
Zhongming Zhao
Huiru Zheng
Bin Zhou
Shuigeng Zhou
Xuezhong Zhou
Daming Zhu
Dongxiao Zhu
Shanfeng Zhu
Xiaoqin Zou
Xiufen Zou
Chiara Zucco

Contents – Part II

Little-Big Data. Reducing the Complexity and Facing Uncertainty of Highly Underdetermined Phenotype Prediction Problems

Know-GRRF: Domain-Knowledge Informed Biomarker Discovery with Random Forests	3
<i>Xin Guan and Li Liu</i>	
Sampling Defective Pathways in Phenotype Prediction Problems via the Fisher’s Ratio Sampler	15
<i>Ana Cernea, Juan Luis Fernández-Martínez, Enrique J. deAndrés-Galiana, Francisco Javier Fernández-Ovies, Zulima Fernández-Muñiz, Oscar Alvarez-Machancoses, Leorey Saligan, and Stephen T. Sonis</i>	
Sampling Defective Pathways in Phenotype Prediction Problems via the Holdout Sampler	24
<i>Juan Luis Fernández-Martínez, Ana Cernea, Enrique J. deAndrés-Galiana, Francisco Javier Fernández-Ovies, Zulima Fernández-Muñiz, Oscar Alvarez-Machancoses, Leorey Saligan, and Stephen T. Sonis</i>	
Comparison of Different Sampling Algorithms for Phenotype Prediction	33
<i>Ana Cernea, Juan Luis Fernández-Martínez, Enrique J. deAndrés-Galiana, Francisco Javier Fernández-Ovies, Zulima Fernández-Muñiz, Óscar Alvarez-Machancoses, Leorey Saligan, and Stephen T. Sonis</i>	

Biomedical Engineering

Composite Piezoelectric Material for Biomedical Micro Hydraulic System . . .	49
<i>Arvydas Palevicius, Giedrius Janusas, Elingas Cekas, and YatinkumarRajeshbhai Patel</i>	
Trabecular Bone Score in Overweight and Normal-Weight Young Women . . .	59
<i>Abdel-Jalil Berro, Marie-Louise Ayoub, Antonio Pinti, Said Ahmaidi, Georges El Khoury, César El Khoury, Eddy Zakhem, Bernard Cortet, and Rawad El Hage</i>	

Sarcopenia and Hip Structure Analysis Variables in a Group of Lebanese Postmenopausal Women	69
<i>Riad Nasr, Eric Watelain, Antonio Pinti, Hayman Saddik, Ghassan Maalouf, Abdel-Jalil Berro, Abir Alwan, César El Khoury, Ibrahim Fayad, and Rawad El Hage</i>	
Feet Fidgeting Detection Based on Accelerometers Using Decision Tree Learning and Gradient Boosting	75
<i>Julien Esseiva, Maurizio Caon, Elena Mugellini, Omar Abou Khaled, and Kamiar Aminian</i>	
Matching Confidence Masks with Experts Annotations for Estimates of Chromosomal Copy Number Alterations	85
<i>Jorge Muñoz-Minjares, Yuriy S. Shmaliy, Tatiana Popova, and R. J. Perez-Chimal</i>	
Using Orientation Sensors to Control a FES System for Upper-Limb Motor Rehabilitation	95
<i>Andrés F. Ruiz-Olaya, Alberto López-Delis, and Adson Ferreira da Rocha</i>	
A Real-Time Research Platform for Intent Pattern Recognition: Implementation, Validation and Application	106
<i>Andres F. Ruiz-Olaya, Gloria M. Díaz, and Alberto López-Delis</i>	
Augmented Visualization and Touchless Interaction with Virtual Organs	118
<i>Lucio Tommaso De Paolis</i>	
Decreased Composite Indices of Femoral Neck Strength in Young Obese Women	128
<i>Abdel-Jalil Berro, Said Ahmaidi, Antonio Pinti, Abir Alwan, Hayman Saddik, Joseph Matta, Fabienne Frenn, Maroun Rizkallah, Ghassan Maalouf, and Rawad El Hage</i>	
On the Use of Decision Trees Based on Diagnosis and Drug Codes for Analyzing Chronic Patients	135
<i>Cristina Soguero-Ruiz, Ana Alberca Díaz-Plaza, Pablo de Miguel Bohoyo, Javier Ramos-López, Manuel Rubio-Sánchez, Alberto Sánchez, and Inmaculada Mora-Jiménez</i>	
Biomedical Image Analysis	
Stochastic Geometry for Automatic Assessment of Ki-67 Index in Breast Cancer Preparations	151
<i>Marek Kowal, Marcin Skobel, Józef Korbicz, and Roman Monczak</i>	

Detection Methods of Static Microscopic Objects	163
<i>Libor Hargaš, Zuzana Loncová, Dušan Koniar, František Jablončík, and Jozef Volák</i>	
Parkinson’s Disease Database Analysis of Stereotactic Coordinates Related to Clinical Outcomes	176
<i>Francisco Estella, Esther Suarez, Beatriz Lozano, Elena Santamarta, Antonio Saiz, Fernando Rojas, Ignacio Rojas, and Fernando Seijo</i>	
Quantitative Ultrasound of Tumor Surrounding Tissue for Enhancement of Breast Cancer Diagnosis	186
<i>Ziemowit Klimonda, Katarzyna Dobruch-Sobczak, Hanna Piotrkowska-Wróblewska, Piotr Karwat, and Jerzy Litniewski</i>	
A Texture Analysis Approach for Spine Metastasis Classification in T1 and T2 MRI	198
<i>Mohamed Amine Larhmam, Saïd Mahmoudi, Stylianos Drisis, and Mohammed Benjelloun</i>	
Parametric Variations of Anisotropic Diffusion and Gaussian High-Pass Filter for NIR Image Preprocessing in Vein Identification	212
<i>Ayca Kirimtat and Ondrej Krejcar</i>	
FLIR vs SEEK in Biomedical Applications of Infrared Thermography.	221
<i>Ayca Kirimtat and Ondrej Krejcar</i>	
Advances in Homotopy Applied to Object Deformation.	231
<i>Jose Alejandro Salazar-Castro, Ana Cristina Umaquinga-Criollo, Lilian Dayana Cruz-Cruz, Luis Omar Alpala-Alpala, Catalina González-Castaño, Miguel A. Becerra-Botero, Diego Hernán Peluffo-Ordóñez, and Cesar Germán Castellanos-Domínguez</i>	
Thermal Imaging for Localization of Anterior Forearm Subcutaneous Veins	243
<i>Orcan Alpar and Ondrej Krejcar</i>	
Detection of Irregular Thermoregulation in Hand Thermography by Fuzzy C-Means	255
<i>Orcan Alpar and Ondrej Krejcar</i>	
Medical Image Classification with Hand-Designed or Machine-Designed Texture Descriptors: A Performance Evaluation	266
<i>Joke A. Badejo, Emmanuel Adetiba, Adekunle Akinrinmade, and Matthew B. Akanle</i>	

Classification of Breast Cancer Histopathological Images Using KAZE Features	276
<i>Daniel Sanchez-Morillo, Jesús González, Marcial García-Rojo, and Julio Ortega</i>	
Biomedical Signal Analysis	
Low Data Fusion Framework Oriented to Information Quality for BCI Systems	289
<i>Miguel Alberto Becerra, Karla C. Alvarez-Urbe, and Diego Hernán Peluffo-Ordoñez</i>	
New Parameter Available in Phonocardiogram for Blood Pressure Estimation	301
<i>Omari Tahar, Ouacif Nadia, Benali Redouane, Dib Nabil, and Bereksi-Reguig Fethi</i>	
Some False ECG Waves Detections Revised by Fractal Dimensions	311
<i>Ibticeme Sedjelmaci and Fethi Bereksi Reguig</i>	
Challenges in Smart and Wearable Sensor Design for Mobile Health	
Reconstruction of Equivalent Electrical Sources on Heart Surface	325
<i>Galina V. Zhikhareva, Mikhail N. Kramm, Oleg N. Bodin, Ralf Seepold, Anton I. Chernikov, Yana A. Kupriyanova, and Natalija A. Zhuravleva</i>	
WearIT - A Rapid Prototyping Platform for Wearables	335
<i>Isabel Leber and Natividad Martínez Madrid</i>	
A Review of Health Monitoring Systems Using Sensors on Bed or Cushion	347
<i>Massimo Conti, Simone Orcioni, Natividad Martínez Madrid, Maksym Gaiduk, and Ralf Seepold</i>	
Textile Sensor Platform (TSP) - Development of a Textile Real-Time Electrocardiogram	359
<i>Thomas Walzer, Christian Thies, Klaus Meier, and Natividad Martínez Madrid</i>	
Sensor-Mesh-Based System with Application on Sleep Study	371
<i>Maksym Gaiduk, Bruno Vunderl, Ralf Seepold, Juan Antonio Ortega, and Thomas Penzel</i>	
Wearable Pneumatic Sensor for Non-invasive Continuous Arterial Blood Pressure Monitoring	383
<i>Viacheslav Antsiperov and Gennady Mansurov</i>	

Healthcare and Diseases

Gene-Gene Interaction Analysis: Correlation, Relative Entropy and Rough Set Theory Based Approach 397
Sujay Saha, Sukriti Roy, Anupam Ghosh, and Kashi Nath Dey

A Transferable Belief Model Decision Support Tool over Complementary Clinical Conditions 409
Abderraouf Hadj Henni, David Pasquier, and Nacim Betrouni

An Online Viewer of FHR Signal for Research, E-Learning and Tele-Medicine 421
Samuel Boudet, Agathe Houzé de l’Aulnoit, Antonio Pinti, Romain Demailly, Michael Genin, Regis Beuscart, Jessica Schiro, Laurent Peyrodie, and Denis Houzé de l’Aulnoit

Modeling Spread of Infectious Diseases at the Arrival Stage of Hajj 430
Sultanah M. Alshammari and Armin R. Mikler

Exploring In-Game Reward Mechanisms in Diaquarium – A Serious Game for Children with Type 1 Diabetes 443
Ida Charlotte Rønningen, Eirik Årsand, and Gunnar Hartvigsen

An “Awareness” Environment for Clinical Decision Support in e-Health 456
Obinna Anya and Hissam Tawfik

Correction to: Textile Sensor Platform (TSP) - Development of a Textile Real-Time Electrocardiogram C1
Thomas Walzer, Christian Thies, Klaus Meier, and Natividad Martínez Madrid

Author Index 469

Contents – Part I

Bioinformatics for Healthcare and Diseases

Trends in Online Biomonitoring	3
<i>Antonín Bárta, Pavel Souček, Vladyslav Bozhynov, Pavla Urbanová, and Dinara Bekkozhaeva</i>	
SARAEasy: A Mobile App for Cerebellar Syndrome Quantification and Characterization	15
<i>Haitham Maarouf, Vanessa López, Maria J. Sobrido, Diego Martínez, and Maria Taboada</i>	
Case-Based Reasoning Systems for Medical Applications with Improved Adaptation and Recovery Stages	26
<i>X. Blanco Valencia, D. Bastidas Torres, C. Piñeros Rodríguez, D. H. Peluffo-Ordóñez, M. A. Becerra, and A. E. Castro-Ospina</i>	

Bioinformatics Tools to Integrate Omics Dataset and Address Biological Question

Constructing a Quantitative Fusion Layer over the Semantic Level for Scalable Inference	41
<i>Andras Gezsi, Bence Bruncsics, Gabor Guta, and Peter Antal</i>	

Challenges and Advances in Measurement and Self-Parametrization of Complex Biological Systems

Effects of External Voltage in the Dynamics of Pancreatic β -Cells: Implications for the Treatment of Diabetes	57
<i>Ramón E. R. González, José Radamés Ferreira da Silva, and Romildo Albuquerque Nogueira</i>	
ISaaC: Identifying Structural Relations in Biological Data with Copula-Based Kernel Dependency Measures	71
<i>Hossam Al Meer, Raghvendra Mall, Ehsan Ullah, Nasreddine Megrez, and Halima Bensmail</i>	
Inspecting the Role of PI3K/AKT Signaling Pathway in Cancer Development Using an <i>In Silico</i> Modeling and Simulation Approach	83
<i>Pedro Pablo González-Pérez and Maura Cárdenas-García</i>	

Cardiac Pulse Modeling Using a Modified van der Pol Oscillator and Genetic Algorithms	96
<i>Fabián M. Lopez-Chamorro, Andrés F. Arciniegas-Mejía, David Esteban Imbajoa-Ruiz, Paul D. Rosero-Montalvo, Pedro García, Andrés Eduardo Castro-Ospina, Antonio Acosta, and Diego Hernán Peluffo-Ordóñez</i>	
Visible Aquaphotomics Spectrophotometry for Aquaculture Systems	107
<i>Vladyslav Bozhynov, Pavel Soucek, Antonin Barta, Pavla Urbanova, and Dinara Bekkozhayeva</i>	
Resolution, Precision, and Entropy as Binning Problem in Mass Spectrometry	118
<i>Jan Urban</i>	
Discrimination Between Normal Driving and Braking Intention from Driver's Brain Signals	129
<i>Efraín Martínez, Luis Guillermo Hernández, and Javier Mauricio Antelis</i>	
Unsupervised Parametrization of Nano-Objects in Electron Microscopy	139
<i>Pavla Urbanová, Norbert Cyran, Pavel Souček, Antonín Bárta, Vladyslav Bozhynov, Dinara Bekkozhayeva, Petr Císař, and Miloš Železný</i>	
Computational Genomics	
Models of Multiple Interactions from Collinear Patterns.	153
<i>Leon Bobrowski and Paweł Zabielski</i>	
Identification of the Treatment Survivability Gene Biomarkers of Breast Cancer Patients via a Tree-Based Approach	166
<i>Ashraf Abou Tabl, Abedalrhman Alkhateeb, Luis Rueda, Waguih ElMaraghy, and Alioune Ngom</i>	
Workflows and Service Discovery: A Mobile Device Approach	177
<i>Ricardo Holthausen, Sergio Díaz-Del-Pino, Esteban Pérez-Wohlfeil, Pablo Rodríguez-Brazzarola, and Oswaldo Trelles</i>	
Chloroplast Genomes Exhibit Eight-Cluster Structuredness and Mirror Symmetry	186
<i>Michael Sadovsky, Maria Senashova, and Andrew Malyshev</i>	
Are Radiosensitive and Regular Response Cells Homogeneous in Their Correlations Between Copy Number State and Surviving Fraction After Irradiation?	197
<i>Joanna Tobiasz, Najla Al-Harbi, Sara Bin Judia, Salma Majid, Ghazi Alsbeih, and Joanna Polanska</i>	

Computational Proteomics

- Protein Tertiary Structure Prediction via SVD and PSO Sampling 211
*Óscar Álvarez, Juan Luis Fernández-Martínez, Ana Cernea,
 Zulima Fernández-Muñiz, and Andrzej Kloczkowski*
- Fighting Fire with Fire: Computational Prediction of Microbial Targets
 for Bacteriocins 221
Edgar D. Coelho, Joel P. Arrais, and José Luís Oliveira
- A Graph-Based Approach for Querying Protein-Ligand Structural Patterns . . . 235
Renzo Angles and Mauricio Arenas

Computational Systems for Modelling Biological Processes

- Predicting Disease Genes from Clinical Single Sample-Based PPI Networks . . . 247
*Ping Luo, Li-Ping Tian, Bolin Chen, Qianghua Xiao,
 and Fang-Xiang Wu*
- Red Blood Cell Model Validation in Dynamic Regime 259
*Kristína Kovalčíková, Alžbeta Bohiniková, Martin Slavík,
 Isabelle Mazza Guimaraes, and Ivan Cimrák*
- Exploiting Ladder Networks for Gene Expression Classification 270
Guray Golcuk, Mustafa Anil Tuncel, and Arif Canakoglu
- Simulation of Blood Flow in Microfluidic Devices for Analysing
 of Video from Real Experiments 279
*Hynek Bachratý, Katarína Bachratá, Michal Chovanec,
 František Kajánek, Monika Smiešková, and Martin Slavík*
- Alignment-Free Z-Curve Genomic Cepstral Coefficients and Machine
 Learning for Classification of Viruses 290
*Emmanuel Adetiba, Oludayo O. Olugbara, Tunmike B. Taiwo,
 Marion O. Adebisi, Joke A. Badejo, Matthew B. Akanle,
 and Victor O. Matthews*
- A Combined Approach of Multiscale Texture Analysis and Interest
 Point/Corner Detectors for Microcalcifications Diagnosis 302
*Liliana Losurdo, Annarita Fanizzi, Teresa M. A. Basile, Roberto Bellotti,
 Ubaldo Bottigli, Rosalba Dentamaro, Vittorio Didonna, Alfonso Fausto,
 Raffaella Massafra, Alfonso Monaco, Marco Moschetta,
 Ondina Popescu, Pasquale Tamborra, Sabina Tangaro,
 and Daniele La Forgia*
- An Empirical Study of Word Sense Disambiguation for Biomedical
 Information Retrieval System 314
Mohammed Rais and Abdelmonaime Lachkar

Drug Delivery System Design Aided by Mathematical Modelling and Experiments

- Modelling the Release of Moxifloxacin from Plasma Grafted Intraocular Lenses with Rotational Symmetric Numerical Framework 329
Kristinn Gudnason, Sven Sigurdsson, Fjola Jonsdottir, A. J. Guiomar, A. P. Vieira, P. Alves, P. Coimbra, and M. H. Gil

Generation, Management and Biological Insights from Big Data

- Predicting Tumor Locations in Prostate Cancer Tissue Using Gene Expression 343
Osama Hamzeh, Abedalrhman Alkhateeb, and Luis Rueda
- Concept of a Module for Physical Security of Material Secured by LIMS 352
Pavel Blazek, Kamil Kuca, and Ondrej Krejcar
- scFeatureFilter: Correlation-Based Feature Filtering for Single-Cell RNAseq. 364
Angeles Arzalluz-Luque, Guillaume Devailly, and Anagha Joshi

High-Throughput Bioinformatic Tools for Medical Genomics

- NearTrans Can Identify Correlated Expression Changes Between Retrotransposons and Surrounding Genes in Human Cancer 373
Rafael Larrosa, Macarena Arroyo, Rocío Bautista, Carmen María López-Rodríguez, and M. Gonzalo Claros
- An Interactive Strategy to Visualize Common Subgraphs in Protein-Ligand Interaction 383
Alexandre V. Fassio, Charles A. Santana, Fabio R. Cerqueira, Carlos H. da Silveira, João P. R. Romanelli, Raquel C. de Melo-Minardi, and Sabrina de A. Silveira
- Meta-Alignment: Combining Sequence Aligners for Better Results 395
Beat Wolf, Pierre Kuonen, and Thomas Dandekar
- Exploiting In-memory Systems for Genomic Data Analysis 405
Zeeshan Ali Shah, Mohamed El-Kalioby, Tariq Faquih, Moustafa Shokrof, Shazia Subhani, Yasser Alnakhli, Hussain Aljafar, Ashiq Anjum, and Mohamed Abouelhoda
- Improving Metagenomic Assemblies Through Data Partitioning: A GC Content Approach 415
Fábio Miranda, Cassio Batista, Artur Silva, Jefferson Morais, Nelson Neto, and Rommel Ramos

Next Generation Sequencing and Sequence Analysis

Quality Assessment of High-Throughput DNA Sequencing Data
via Range Analysis 429
Ali Fotouhi, Mina Majidi, and M. Oğuzhan Külekci

A BLAS-Based Algorithm for Finding Position Weight Matrix Occurrences
in DNA Sequences on CPUs and GPUs. 439
Jan Fostier

Analyzing the Differences Between Reads and Contigs When Performing
a Taxonomic Assignment Comparison in Metagenomics 450
*Pablo Rodríguez-Brazzarola, Esteban Pérez-Wohlfeil,
Sergio Díaz-del-Pino, Ricardo Holthausen, and Oswaldo Trelles*

Estimating the Length Distributions of Genomic Micro-satellites
from Next Generation Sequencing Data 461
*Xuan Feng, Huan Hu, Zhongmeng Zhao, Xuanping Zhang,
and Jiayin Wang*

CIGenotyper: A Machine Learning Approach for Genotyping
Complex Indel Calls 473
*Tian Zheng, Yang Li, Yu Geng, Zhongmeng Zhao, Xuanping Zhang,
Xiao Xiao, and Jiayin Wang*

Genomic Solutions to Hospital-Acquired Bacterial Infection Identification . . . 486
Max H. Garzon and Duy T. Pham

Interpretable Models in Biomedicine and Bioinformatics

Kernel Conditional Embeddings for Associating Omic Data Types 501
Ferran Reverter, Esteban Vegas, and Josep M. Oller

Metastasis of Cutaneous Melanoma: Risk Factors, Detection
and Forecasting. 511
*Iker Malaina, Leire Legarreta, Maria Dolores Boyano,
Jesus Gardeazabal, Carlos Bringas, Luis Martinez,
and Ildefonso Martinez de la Fuente*

Graph Theory Based Classification of Brain Connectivity Network
for Autism Spectrum Disorder 520
Ertan Tolan and Zerrin Isik

Detect and Predict Melanoma Utilizing TCBR and Classification
of Skin Lesions in a Learning Assistant System 531
Sara Nasiri, Matthias Jung, Julien Helsper, and Madjid Fathi

On the Use of Betweenness Centrality for Selection of Plausible Trajectories in Qualitative Biological Regulatory Networks	543
<i>Muhammad Tariq Saeed, Jamil Ahmad, and Amjad Ali</i>	

Author Index	553
-------------------------------	-----

**Little-Big Data. Reducing the
Complexity and Facing Uncertainty of
Highly Underdetermined Phenotype
Prediction Problems**



Know-GRRF: Domain-Knowledge Informed Biomarker Discovery with Random Forests

Xin Guan^{1,2,3(✉)} and Li Liu^{2,3,4(✉)}

¹ Intel Corporation, Chandler, AZ 85226, USA

² Department of Biomedical Informatics, Scottsdale, AZ 85259, USA

³ Biodesign Institute, Arizona State University, Tempe, AZ 85287, USA
{xinguan, liliu}@asu.edu

⁴ Department of Neurology, Mayo Clinic, Scottsdale, AZ 85259, USA

Abstract. Due to its robustness and built-in feature selection capability, random forest is frequently employed in omics studies for biomarker discovery and predictive modeling. However, random forest assumes equal importance of all features, while in reality domain knowledge may justify the prioritization of more relevant features. Furthermore, it has been shown that an antecedent feature selection step can improve the performance of random forest by reducing noises and search space. In this paper, we present a novel Know-guided regularized random forest (Know-GRRF) method that incorporates domain knowledge in a random forest framework for feature selection. Via rigorous simulations, we show that Know-GRRF outperforms existing methods by correctly identifying informative features and improving the accuracy of subsequent predictive models. Know-GRRF is responsive to a wide range of tuning parameters that help to better differentiate candidate features. Know-GRRF is also stable from run to run, making it robust to noises. We further proved that Know-GRRF is a generalized form of existing methods, RRF and GRRF. We applied Know-GRRF to a real world radiation biodosimetry study that uses non-human primate data to discover biomarkers for human applications. By using cross-species correlation as domain knowledge, Know-GRRF was able to identify three gene markers that significantly improved the cross-species prediction accuracy. We implemented Know-GRRF as an R package that is available through the CRAN archive.

Keywords: Biomarker discovery · Domain knowledge · Feature selection
Regularized random forest

1 Introduction

In biomarker discovery, domain knowledge offers valuable information on biological relevance of candidate features. This is especially important in analyzing omics data, in which a large number of molecular features are noises and may cause over-fitting. Furthermore, because domain knowledge can help identify functionally impactful biomarkers, the interpretation of the selected biomarkers and predictive model can provide useful insight into the underlying biological mechanisms.

Computationally, the idea of incorporating domain knowledge in feature selection has been studied widely. The most straightforward and commonly used approach is to apply domain knowledge as a “pre-processing” filter or “ad-hoc” explanation after modeling. In these cases, candidate features are ranked according to a set of rules and annotations, and top-ranked features are selected [1, 2]. Approaches that are more sophisticated seek systematic ways to quantify and optimize the gain of performance via incorporation of domain knowledge. For example, Barzilay studied SVM and constructed a kernel function with domain knowledge in texture recognition [3]. Ding utilized mutual information as a measure of relevance of genes and proposed a minimum redundancy maximum relevance feature selection framework [4]. More recently, an interaction-based feature-selection method was proposed based on random lasso that imposes strong penalty on genes corresponding to their dependency [5].

In this study, we reported an embedded method to incorporate domain knowledge for feature selection in the framework of random forest (RF). RF is a powerful supervised learning method with a built-in feature selection capability. It has been widely used in high-dimensional data analysis, such as biomarker discovery from omics data. Several studies reported that an antecedent feature selection step can significantly improve the performance of RF [6, 7]. Among many feature selection algorithms, regularized RF (RFF) and guided RFF (GRFF) [8] are two popular methods that reside in the same framework of RF. Unfortunately, feature sets selected by RFF and GRFF show a large variation due to the randomness involved in the subsampling step and the node-splitting step. The selection of tuning parameters of these methods is highly subjective and the performance responds to only a narrow range of parameters. Furthermore, these tuning parameters are computed based on information intrinsic to the RF algorithm, thus unable to incorporate external domain knowledge. To address these problems, we developed a Know-GRRF method. In various simulated scenarios, we demonstrated that Know-GRRF could identify truly informative and relevant features more accurately than RFF and GRFF. We also proved that Know-GRRF is a generalized form of RFF and GRFF and provides higher sensitivity to tuning parameters and higher differentiability of similar features than these methods.

For a real-world application, we analyzed a proteomic dataset that aimed to identify human biomarkers for radiation biodosimetry using non-human primate as experimental subjects. By using Know-GRRF with cross-species correlation as prior knowledge, we identified three gene markers and built a predictive model that significantly outperformed the original 28-gene model. We implement Know-GRRF as an R package that is available in the CRAN archive.

2 New Method

2.1 Know-GRRF Algorithm

The ordinary RF algorithm is a supervised learning method with a built-in mechanism to perform feature selection at each splitting node [9]. Specifically, given a data set consisting of N observations, each measured on a response variable Y and P predictor variables (i.e. features X_i where $i = 1, \dots, P$), RF learns multiple decision trees to

model the relationship between X and Y in bootstrapped samples. At each splitting node v of a tree, the feature with the highest information gain $Gain(X_i, v)$ is selected. Know-GRRF takes advantage of this mechanism by defining a penalty coefficient λ_i that is an exponential function of $score_i$ derived from domain knowledge.

$$\lambda_i = score_i^\delta \quad (1)$$

where $score_i \in [0, 1]$, δ is the tuning parameter and λ is constrained between 0 and 1. We use λ_i to regularize $Gain(X_i, v)$ as

$$Gain_R(X_i, v) = \begin{cases} \lambda_i Gain(X_i, v) & i \notin F \\ Gain(X_i, v) & i \in F \end{cases} \quad (2)$$

where F is a set of features selected in previous nodes. F begins as an empty set and accumulates selected features used for splitting at each node. If feature i is not selected in previous nodes, λ_i penalizes it by requesting higher information gain and stronger relevant based on prior knowledge. A smaller λ leads to a larger penalty. The feature with the highest $Gain_R(X_i, v)$ is then added to the feature set F . As such, Know-GRRF eliminates redundant and/or irrelevant features effectively.

The objective of Known-GRRF is to select a compact feature subset F without loss of predictive information about Y . We use Akaike's information criterion (AIC) of out-of-bag (OOB) predictions as the loss function.

$$AIC = 2k - 2 \ln(\hat{L}) \quad (3)$$

where k is the number of features in F and $\ln(\hat{L})$ is a goodness-of-fit measure. We define $\ln(\hat{L})$ for a classification task and for a regression task as

$$\ln(\hat{L}) = \begin{cases} \sum_1^N Y_j \ln \hat{p}_j + (1 - Y_j) \ln(1 - \hat{p}_j) & (classification) \\ -\frac{N}{2} \ln \frac{\sum_1^S \sum_N (y_j - \hat{y}_j)^2}{N} & (regression) \end{cases} \quad (4)$$

In classification problems, \hat{p}_j is the probability of the j^{th} sample being predicted as Y_j . If $\hat{p}_j = 0$, it is replaced by $\hat{p}_j = \frac{1}{2N}$, and if $\hat{p}_j = 1$, it is replaced by $\hat{p}_j = 1 - \frac{1}{2N}$. We then optimize the tuning parameter δ within a user-specified range to minimize AIC.

In practice, a big challenge is the large variance from run to run due to the randomness of bootstrapping. To address this problem, after obtaining the optimal δ value, we perform a stability selection. Specifically, with the optimal δ value, we build Known-GRRF models m times that return m sets of selected features. We then add features sequentially according to the selection frequency and select the set of features that lead to the smallest AIC value.

2.2 Relationship with RFF and GRFF

Equation (2) that utilizes penalized information gain for regularization lies in the center of RRF, GRFF and Know-GRRF. In RFF, the penalty coefficient λ_i is the variable importance (VI_i) score that is computed by the ordinary RF at each node. In GRRF, λ_i is defined as a function of normalized VI (VI'_i)

$$\lambda_i = (1 - \gamma) + \gamma VI'_i \quad (5)$$

where parameter γ is tuned to optimize the model performance. When $\gamma = 0$, GRRF is equivalent to RRF. In Know-GRRF, λ_i is an exponential function as in Eq. (1). When $\delta = 0$ and $score_i$ taking the value of VI'_i , Know-GRRF is equivalent to GRRF. Therefore, Known-GRRF is a generalized form of RRF and GRFF.

3 Results

3.1 Simulation

Dataset. We compared the performance of Know-GRRF with other RF-based feature selection methods using simulated datasets. We simulated three scenarios with varying levels of complexity, including linear relationship, higher-order relationship and interaction (Table 1). The original values of the response variable (Y) was used in regression tasks, and was dichotomized to binary for classification tasks.

Table 1. True relationship in simulated scenario

Scenario	Relationships
1. Linear	$Y = 0.3X_1 + 0.5X_2 + 0.7X_3 + 0.9X_4 + 1.1X_5 + 1.3X_6 + 1.5X_7 + 1.7X_8 + 1.9X_9 + 2.1X_{10}$
2. Higher order	$Y = 0.9X_4 + 1.1X_5 + 1.3X_6 + 1.7X_8 + 1.9X_9 + 2.1X_{10} + 1.7X_{11}^2$
3. Interaction	$Y = 0.1 + 0.9X_4 + 1.1X_5 + 1.3X_6 + 1.7X_8 + 1.9X_9 + 2.1X_{10} + 1.7X_{11}X_{12}$

For each scenario, we generated 100 *i.i.d.* random variables (X_1, \dots, X_{100}) under normal distribution with 200 observations. We used a half of the observations as the training set for feature selection. To quantify the similarity between two feature sets (i.e. simulated explanatory set F_1 vs. method selected set F_2), we computed the Jaccard Index ($JI = |F_1 \cap F_2| / |F_1 \cup F_2|$), true positive rate ($TPR = |F_1 \cap F_2| / |F_1|$) and false positive rate ($FPR = |F_2 - F_1| / |F_2|$). We then applied the selected features to the remaining half of the observations for predictive modeling using the ordinary RF with default settings. To account for variability, we built ten RF models with each feature set and reported the average performance. Because there is no domain knowledge in simulated data, we used intrinsic data characteristics as surrogates.

Classification tasks. To find surrogates of domain knowledge, we performed two-group t-test followed by Benjamini-Hochberg correction to derive the q-value [10].

We normalized the q -value as $q'_i = \frac{1-q_i}{\max_{j=1}^p(1-q_j)}$ to constrain q'_i within $[0, 1]$ and used it as domain knowledge $score_i$ for Know-GRRF. We set the range of δ as $[0.1, 10]$. For comparison, we tested R/varSelRF that implements RRF for classification with default settings and tested GRRF with $\gamma = 0.5$ for optimal performance. As a negative control, we built RF models with no antecedent feature selection. To evaluate the performance, we computed mean standard error (MSE) of OOB samples (Table 2).

Table 2. Methods comparison in two-class classification tasks

Method	Selected features	JI	TPR	FPR	OOB MSE (SD)
Scenario 1					
Ground truth	$X_1, \dots, 10$	–	–	–	0.19 (0.016)
No feature selection	$X_1, \dots, 100$	0.10	100%	90%	0.30 (0.036)
varSelRF	$X_5, 9, 28, 36, 42$	0.15	20%	60%	0.28 (0.011)
GRRF ($\gamma = 0.5$)	$X_5, 6, 7, 9, 12, 28, 42, 61$	0.29	40%	50%	0.24 (0.020)
Know-GRRF ($\delta = 2.30$, Stability = 100%)	$X_5, 6, 7, 9, 10, 42$	0.45	50%	17%	0.25 (0.012)
Scenario 2					
Ground truth	$X_{4, 5, 6, 8, 9, 10, 11}$	–	–	–	0.22 (0.019)
No feature selection	$X_1, \dots, 100$	0.07	100%	93%	0.30 (0.012)
varSelRF	$X_{4, 5, 9, 10, 30, 61}$	0.44	57%	33%	0.26 (0.014)
GRRF ($\gamma = 0.5$)	$X_{4, 5, 9, 10, 61, 85}$	0.44	57%	33%	0.27 (0.015)
Know-GRRF $\delta = 1.97$, Stability = 60%)	$X_{4, 5, 6, 9, 10, 61}$	0.63	71%	17%	0.23 (0.020)
Scenario 3					
Ground truth	$X_{4, 5, 6, 8, 9, 10, 11, 12}$	–	–	–	0.24 (0.018)
No feature selection	$X_1, \dots, 100$	0.08	100%	92%	0.31 (0.025)
varSelRF	$X_5, 8, 9, 10, 29$	0.44	50%	20%	0.24 (0.012)
GRRF ($\gamma = 0.5$)	$X_5, 8, 10$	0.38	38%	0%	0.27 (0.014)
Know-GRRF ($\delta = 0.81$, Stability = 100%)	$X_5, 8, 9, 10, 52$	0.44	50%	20%	0.25 (0.020)

In all three scenarios, Know-GRRF reported the largest JI values among the three tested methods, supporting its highest similarities to the ground truth. It also has the highest TPRs, supporting its capability to recover the most informative and relevant features. Except in scenario 3, Know-GRRF also reported the lowest FPRs. OOB error of RF models using feature selected by Know-GRRF was comparable or better than the other two methods. Interestingly, in scenario 3 that simulated higher-order relationship between features, no method was able to select X_{11} or X_{12} . The reason may be because X_{11} or X_{12} by itself had no effect. Thus, the q -values and VI were insignificant for these two features. In fact, such scenarios where interaction is significant at the absence of main effects has been a challenging task for all methods [11].

Regression tasks. We used VI'_i as domain knowledge *score* _{i} in Know-GRRF and set the range of δ as [0.1, 10]. For comparisons, we tested R/VSURF that implements RRF for regression with default settings and tested GRRF with $\gamma = 0.9$ for optimal performance. We summarized the performance of these methods in Table 3.

Table 3. Methods Comparison in Regression Tasks

Method	Selected features	JI	TPR	FPR	OOB MSE (SD)
Scenario 1					
Ground truth	$X_1, \dots, 10$	–	–	–	0.08 (0.022)
No feature selection	$X_1, \dots, 100$	0.10	100%	90%	0.14 (0.024)
varSelRF	$X_6, 7, 8, 9, 10$	0.50	50%	0%	0.08 (0.016)
GRRF ($\gamma = 0.5$)	$X_3, 4, 6, 7, 8, 10$	0.60	60%	0%	0.10 (0.016)
Know-GRRF ($\delta = 2.30$, Stability = 100%)	$X_3, 4, 6, 7, 8, 9, 10$	0.70	70%	0%	0.08 (0.015)
Scenario 2					
Ground truth	$X_4, 5, 6, 8, 9, 10, 11$	–	–	–	0.36 (0.047)
No feature selection	$X_1, \dots, 100$	0.07	100%	93%	0.33 (0.048)
varSelRF	$X_8, 10, 11$	0.43	43%	0%	0.33 (0.043)
GRRF ($\gamma = 0.5$)	$X_6, 8, 9, 10, 11, 41, 51$	0.56	71%	29%	0.36 (0.033)
Know-GRRF ($\delta = 1.97$, Stability = 60%)	$X_8, 10, 11$	0.43	43%	0%	0.33 (0.033)
Scenario 3					
Ground truth	$X_4, 5, 6, 8, 9, 10, 11, 12$	–	–	–	0.19 (0.016)
No feature selection	$X_1, \dots, 100$	0.08	100%	92%	0.18 (0.021)
varSelRF	$X_6, 8, 9, 10$	0.50	50%	0%	0.19 (0.026)
GRRF ($\gamma = 0.5$)	$X_6, 8, 9, 10, 51, 93$	0.40	50%	33%	0.19 (0.020)
Know-GRRF ($\delta = 0.81$, Stability = 100%)	$X_6, 8, 9, 10$	0.50	50%	0%	0.19 (0.017)

For regression problems, performance of Know-GRRF and VSURF are quite comparable, but better than GRRF consistently. They selected fewer features and resulted in smaller MSE than GRRF. In scenario 1, Know-GRRF reported the highest JI value and TPR value and the lowest FPR value among the three tested methods, supporting its highest similarities to the ground truth. In scenario 3, Know-GRRF performed as well as RFF and both exceeded GRRF. In scenario 2, GRRF performed the best and Know-GRRF and RFF performed similarly.

3.2 Radiation Data

Dataset. We applied Know-GRRF to data from a radiation biodosimetry study [12]. The goal of this study is to discover gene expression biomarkers and develop a biodosimetry model to estimate absorbed dose by humans in an ionizing radiation

exposure event. Initially, researchers used non-human primates (NHPs) as test subjects. In a case-control design, NHPs were exposed to varying levels of radiation ranging from 0 Gy to ~ 10 Gy. Biomarkers and biodosimetry models obtained from NHPs were validated in human subjects who went through radiation as part of cancer therapy (Table 4). As described in the original study, the gene expression profile of each NHP test subject and human participant was measured using protein microarrays. The final biodosimetry model consisted of 28 genes. When evaluated on human subjects, the MSE was 7.82.

Table 4. Model comparison using different sets of features

Methods	Number of genes	MSE (SD)	
		NHP	Human
Original	28	1.52 (0.025)	7.82 (0.058)
VSURF	5	1.62 (0.017)	9.51 (0.160)
GRRF (VI, $\gamma = 0.9$)	12	1.59 (0.019)	8.30 (0.074)
GRRF (CC, $\gamma = 0.9$)	25	1.52 (0.026)	7.31 (0.064)
Know-GRRF ($\delta = 12$, stability = 100%)	3	3.21 (0.075)	4.01 (0.050)

Performance. Because this study transferred a model learned from NHPs to human subjects, we applied the Know-GRRF with domain knowledge of cross-species correlation. Specifically, we first mapped each of the 28 NHP biomarker gene to a unique human gene based on gene orthologs. Then, for each pair i , we retrieved the cross-species correlation coefficient (CC_i) reported in the original study. This domain knowledge was translated into $score_i = \frac{CC_i}{\max_{i=1}^p CC_i}$ so that $score_i \in [0, 1]$. We then performed feature selection on NHP data using Know-GRRF with $\delta \in [0.1, 20]$ followed by stability selection with 10 runs. In each stability run, a RF model was trained using features selected from NHP data and applied to human data to predict radiation dose. Know-GRRF reported an optimal δ value of 12 and stability of 100%. With these criteria, Know-GRRF identified three genes (*CDKN1A*, *CXXC5* and *MYC*) as informative biomarkers. The MSE of the predictive model was reduced to 4.01 in human. Compared to the original model consisting of 28 genes, this new model reduced the complexity by 9.3 times and reduced the predictive error by 49%. For comparison, we applied VSURF and GRRF to the same data. We used the default setting for VSURF and two settings for GRRF (one as default and the other with *CC* replacing *I*). With selected features, we built 10 RF models and reported the average MSE (Table 4). Both VSURF and GRRF downsized the selected feature set, although the RF models had similar or worse MSE than the original model. Know-GRRF was the only method that achieved a significantly better performance.

One concern was that Know-GRRF increased MSE on NHP data. A close examination of the three genetic biomarkers identified by Know-GRRF revealed that only *CDKN1A* was among the most dose-responsive genes in NHPs. *CXXC5* and *MYC* were moderately responsive in NHPs. However, *CXXC5* and *MYC* showed high correlation

between NHPs and human. Therefore, Know-GRRF made a compromise to choose features that produced accurate predictions in humans. Interestingly, GRRF that used *CC* to replace *VI* did not achieve a higher accuracy in humans, implicating its inability to incorporate priori knowledge during regularization.

3.3 Tuning Parameters

Know-GRRF and GRRF both use a tuning parameter to adjust the degree of regularization. To examine the influence of the tuning parameter values on feature selection, we varied the values of regularization coefficients δ and γ in Know-GRRF and GRRF, respectively. As expected, the number of selected features decreased with large values of δ or γ (Fig. 1A). However, the regularization effect, i.e. change of the number of selected features was less sensitive to γ as compared to δ . Even when γ reached the maximum value of 1, the size of the final feature set was still relatively large (~ 24). On the other hand, the number of features decreased fast with δ . Consequently, the performance of RF models using the selected feature was more responsive to δ than to γ (Fig. 1B). During optimization, a convex curve of MSE is a strong indicator that an optimal performance is achieved at the lowest point. We observed such a curve in Know-GRRF but not in GRRF. Thus, the confidence associated with the selected features was significantly higher in Know-GRRF than in GRRF.

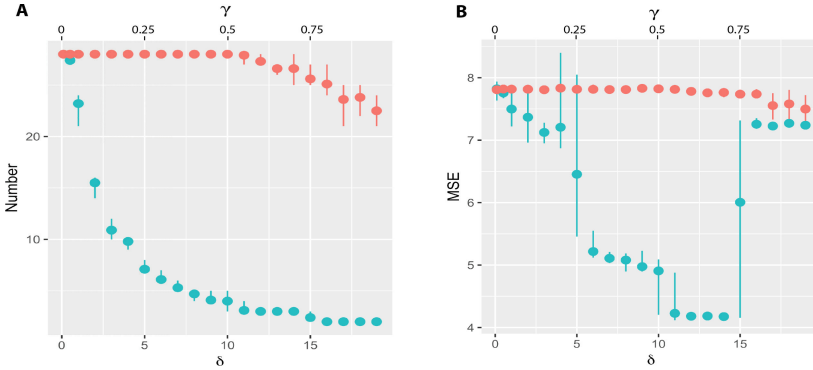


Fig. 1. Comparison of Know-GRRF with GRRF on the number of genes selected (A) and on the MSE in human samples (B). X-axes are tuned parameters δ in Know-GRRF (blue dots) and γ in GRRF (red dots). Y-axes are the number of selected features (A) and the MSE in human samples. Errors bar represent the standard deviation of 10 runs. (Color Figure online)

As a confirmation, we examined the penalty coefficient λ_i for the 28 biomarker genes reported in the original study. We found that λ_i computed in Know-GRRF with $\delta = 0.2$ were well separated from λ_i computed with $\delta = 5$ (Fig. 2). These λ_i values spanned a large spectrum in the allowed range between 0 and 1 and clustered at the two ends. Contrarily, λ_i computed in GRRF with $\gamma = 0.2$ were close to λ_i computed with $\gamma = 0.8$, and concentrated in the range between 0.5 and 1. Therefore, Know-GRRF can better differentiate features than GRRF.

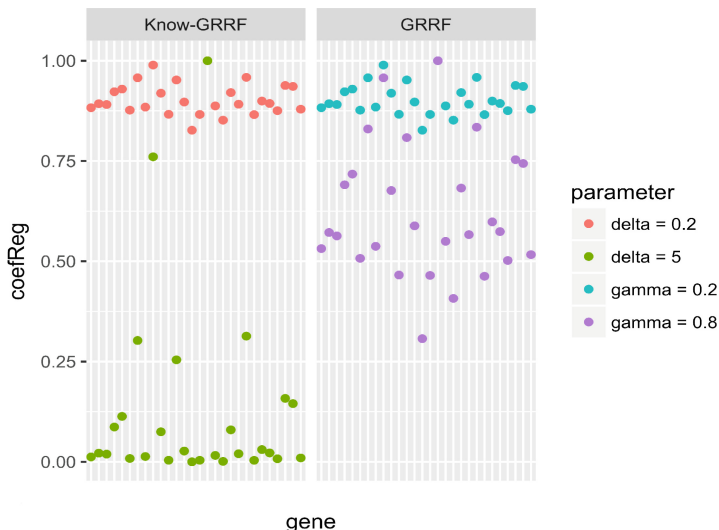


Fig. 2. Penalty coefficient for each of the 28 genes reported in the original study of radiation biodosimetry. The left panel is λ_i in Know-GRRF, at the setting of $\delta = 0.2$ and $\delta = 5$. The right panel is λ_i in GRRF, at the setting of $\gamma = 0.2$ and $\gamma = 0.8$.

Our observations on GRRF were consistent with previous reports by Deng *et al.* who noticed that GRRF was less sensitive to γ when $\gamma > 0.5$ [8]. Because γ is constrained from 0 to 1 and could reach a maximum, it greatly limits the tuning options. On the other hand, δ in Know-GRRF has no upper limit, thus provides users a better control of the degree of regularization.

4 Discussion and Conclusion

In high-dimensional data, the number of observation is always much smaller than the number of features. This curse of dimensionality causes many features to share similar information gain. When an RF algorithm learns a model, it is highly likely to select a feature that is not relevant to the classification or regression problem but is associated with another relevant feature by random chance. In these cases, prior knowledge on the relevance of the candidate features can greatly help eliminate impertinent features and select the truly impactful ones.

In this paper, we reported a new method, Know-GRRF that incorporates priori knowledge in the RF framework for feature selection. Through simulations and applications on a real world biomarker discovery dataset, we demonstrated that Know-GRRF was more robust than existing methods, namely RRF and GRRF. GRRF has been known to be insensitive to the tuning parameter. The Know-GRRF solved this problem by employing an exponential function in the computation of the regularization coefficient. We showed that Know-GRRF was more responsive to the tuned parameter

that is not bound by an upper limit, and is more differentiable on candidate features. As a result, a global minimum can often be found by Know-GRRF following a convex optimization.

Because RF and its derived methods used bootstraps, randomness and variation between runs are unavoidable. Know-GRRF added a stability selection step to find consistently selected features. By running a stability test, Know-GRRF reduced the variability and improve the reproducibility from run to run.

Furthermore, we proved that Know-GRRF is a generalized form of RRF and GRRF. Accuracy of Know-GRRF is comparable to existing implementations of RRF, i.e. VSURF or varSelRF, if not better. However, the computational time of Know-GRRF is much shorter, especially when compared to VSURF that adopted a stepwise selection procedure. We also tested using q -value instead of VI as the weight for each feature in classification. The results suggest the use of other statistics can be successfully applied to the Know-GRRF framework. Similarly, we may use other domain knowledge, such as mutual information and evolutionary weight, to regulate the penalty coefficient.

When we applied Know-GRRF to a radiation biodosimetry study that used NHP data to derive biomarkers for humans, we used cross-species correlation as the domain knowledge to guide the feature selection. Know-GRRF selected only 3 genes from among the original 28 biomarker genes, while cutting the predictive errors by almost 50%. The three biomarker genes are *CDKN1A*, *CXXC5* and *MYC*. *CDKN1A* is a well-known gene to be transcriptionally regulated by p53 in response to stresses such as ionizing radiation [12] and reported by multiple groups as the top candidate gene for use in biodosimetry [13–16]. *CXXC5* is related to DNA damage and was reported as a potential therapeutic target in primary acute myeloid leukemia cells in response to radiation therapy [17]. *MYC* is a novel biomarker for radiation. It is an important gene that regulates a range of cellular processes including biogenesis and protein synthesis [18]. The role of *CXXC5* and *MYC* in response to radiation is worth studying in the future.

There are some limitations of Know-GRRF. First, in our simulation of scenarios with interaction terms, none of those tested methods were able to identify the features with interaction. Because Know-GRRF is not intended to solve XOR problems, we may need some other methods to tackle this challenge. If we have domain knowledge of interaction, for example, gene-gene interaction, we may be able to incorporate it to find interacting features. But this is beyond of the scope of this dissertation. Second, it requires some computation for optimization and build more ensembles than GRRF. To search for the best value of the regularization term, we need to solve a convex optimization problem. We may set a larger tolerance to achieve faster computation, because there is a wide range of minima as solutions. Third, stability test requires more ensembles but can reduce variability from run to run. Although we can assess the variation in the process of stability test, Know-GRRF is still based on RF. Thus, variation is hard to be eliminated. However, we are able to remove random selected features by using a stability test.

In summary, our new method, Know-GRRF is a powerful method to incorporate domain knowledge for feature selection. It can be implemented through R package RRF that is available through the CRAN website.

Acknowledgments. We thank George Runger, Kristin Gillis, Vel Murugan, Jin Park and Garrick Wallstrom for insightful discussions. This project has been funded in part with federal funds from the Biomedical Advanced Research and Development Authority, office of the Assistant Secretary for Preparedness and Response, Office of the Secretary, Department of Health and Human Services under Contract No. HHS01201000008C.

References

1. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., et al.: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999). <https://doi.org/10.1126/science.286.5439.531>
2. Zhou, H., Skolnick, J.: A knowledge-based approach for predicting gene–disease associations. *Bioinformatics* **32**, 2831–2838 (2016). <https://doi.org/10.1093/bioinformatics/btw358>
3. Barzilay, O., Brailovsky, V.L.: On domain knowledge and feature selection using a support vector machine. *Pattern Recognit. Lett.* **20**, 475–484 (1999). [https://doi.org/10.1016/S0167-8655\(99\)00014-8](https://doi.org/10.1016/S0167-8655(99)00014-8)
4. Ding, C., Peng, H.: Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **03**, 185–205 (2005). <https://doi.org/10.1142/S0219720005001004>
5. Park, H., Niida, A., Imoto, S., Miyano, S.: Interaction-based feature selection for uncovering cancer driver genes through copy number-driven expression level. *J. Comput. Biol.* **24**, 138–152 (2017). <https://doi.org/10.1089/cmb.2016.0140>
6. Iguyon, I., Elisseeff, A.: An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3**, 1157–1182 (2003)
7. Bolón-Canedo, V., Sánchez-Marño, N., Alonso-Betanzos, A., Benítez, J.M., Herrera, F.: A review of microarray datasets and applied feature selection methods. *Inf. Sci.* **282**, 111–135 (2014). <https://doi.org/10.1016/j.ins.2014.05.042>
8. Deng, H., Runger, G.: Gene selection with guided regularized random forest. *Pattern Recognit.* **46**, 3483–3489 (2013). <https://doi.org/10.1016/j.patcog.2013.05.018>
9. Breiman, L.: *Classification and Regression Trees*. Wadsworth International Group, Belmont (1984)
10. Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995)
11. Bolón-Canedo, V., Sánchez-Marño, N., Alonso-Betanzos, A.: A review of feature selection methods on synthetic data. *Knowl. Inf. Sys.* **34**, 483–519 (2013). <https://doi.org/10.1007/s10115-012-0487-8>
12. Park, J.G., Paul, S., Briones, N., Zeng, J., Gillis, K., et al.: Developing human radiation biodosimetry models: testing cross-species conversion approaches using an ex vivo model system. *Radiat. Res.* **187**, 708–721 (2017). <https://doi.org/10.1667/RR14655.1>
13. Marchetti, F., Coleman, M.A., Jones, I.M., Wyrobek, A.J.: Candidate protein biodosimeters of human exposure to ionizing radiation. *Int. J. Radiat. Biol.* **82**, 605–639 (2006). <https://doi.org/10.1080/09553000600930103>
14. Paul, S., Barker, C.A., Turner, H.C., McLane, A., Wolden, S.L., et al.: Prediction of in vivo radiation dose status in radiotherapy patients using ex vivo and in vivo gene expression signatures. *Radiat. Res.* **175**, 257–265 (2011). <https://doi.org/10.1667/rr2420.1>

15. Tucker, J.D., Joiner, M.C., Thomas, R.A., Grever, W.E., Bakhmutsky, M.V., et al.: Accurate gene expression-based biodosimetry using a minimal set of human gene transcripts. *Int. J. Radiat. Oncol. Biol. Phys.* **88**, 933–939 (2014). <https://doi.org/10.1016/j.ijrobp.2013.11.248>
16. Riecke, A., Rufa, C.G., Cordes, M., Hartmann, J., Meineke, V., et al.: Gene expression comparisons performed for biodosimetry purposes on in vitro peripheral blood cellular subsets and irradiated individuals. *Radiat. Res.* **178**, 234–243 (2012). <https://doi.org/10.1667/rr2738.1>
17. Bruserud, O., Reikvam, H., Fredly, H., Skavland, J., Hagen, K.M., et al.: Expression of the potential therapeutic target CXXC5 in primary acute myeloid leukemia cells - high expression is associated with adverse prognosis as well as altered intracellular signaling and transcriptional regulation. *Oncotarget* **6**, 2794–2811 (2015). <https://doi.org/10.18632/oncotarget.3056>
18. van Riggelen, J., Yetil, A., Felsher, D.W.: MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev. Cancer* **10**, 301–309 (2010). <https://doi.org/10.1038/nrc2819>



Sampling Defective Pathways in Phenotype Prediction Problems via the Fisher's Ratio Sampler

Ana Cernea¹, Juan Luis Fernández-Martínez¹(✉),
Enrique J. deAndrés-Galiana^{1,2}, Francisco Javier Fernández-Ovies¹,
Zulima Fernández-Muñiz¹, Oscar Alvarez-Machancoses¹,
Leorey Saligan³, and Stephen T. Sonis^{4,5}

¹ Group of Inverse Problems, Optimization and Machine Learning,
Department of Mathematics, University of Oviedo, C/Federico García-Lorca, 18,
33007 Oviedo, Spain
jlfm@uniovi.es

² Department of Informatics and Computer Science,
University of Oviedo, Oviedo, Spain

³ National Institute of Nursing Research, National Institutes of Health,
Bethesda, MD, USA

⁴ Primary Endpoint Solutions, Watertown, MA, USA

⁵ Brigham and Womens' Hospital and the Dana-Farber Cancer Institute,
Boston, MA, USA

Abstract. In this paper, we introduce the Fisher's ratio sampler that serves to unravel the defective pathways in highly underdetermined phenotype prediction problems. This sampling algorithm first selects the most discriminatory genes, that are at the same time differentially expressed, and samples the high discriminatory genetic networks with a prior probability that it is proportional to their individual Fisher's ratio. The number of genes of the different networks is randomly established taking into account the length of the minimum-scale signature of the phenotype prediction problem which is the one that contains the most discriminatory genes with the maximum predictive power. The likelihood of the different networks is established via leave-one-out-cross-validation. Finally, the posterior analysis of the most frequently sampled genes serves to establish the defective biological pathways. This novel sampling algorithm is much faster and simpler than Bayesian Networks. We show its application to a microarray dataset concerning a type of breast cancers with very bad prognosis (TNBC). In these kind of cancers, the breast cancer cells have tested negative for hormone epidermal growth factor receptor 2 (HER-2), estrogen receptors (ER), and progesterone receptors (PR). This lack causes that common treatments like hormone therapy and drugs that target estrogen, progesterone, and HER-2 are ineffective. We believe that the genetic pathways that are identified via the Fisher's ratio sampler, which are mainly related to signaling pathways, provide new insights about the molecular mechanisms that are involved in this complex disease. The Fisher's ratio sampler can be also applied to the genetic analysis of other complex diseases.

1 Introduction

Genomic phenotype prediction is predicated on finding a set of genes that prospectively distinguishes a given phenotype. This kind of problem has a high underdetermined character since the number of monitored genetic probes markedly exceeds the number of collected samples (patients). This imbalance creates uncertainty as, potentially, there exist many equivalent genetic networks that could predict the phenotype with a similar accuracy [1], a feature that creates ambiguity in the characterization of the biological pathways since many genes are highly discriminatory. In top of that, a significant obstacle in the analysis of genetic data is the absence of a conceptual model that relates the different genes/probes to the class prediction between the set of genetic signatures \mathbf{g} and the set of classes $C = \{1, 2\}$ in which the phenotype is divided:

$$\mathbf{L}^*(\mathbf{g}) : \mathbf{g} \in \mathbb{R}^s \rightarrow C = \{1, 2\}. \quad (1)$$

Finding the discriminatory genetic signatures corresponding to $\mathbf{L}^*(\mathbf{g})$ can be interpreted as a generalized regression model. For that purpose, the modelling is dividing in 2 steps: learning and validation. The learning process consists in giving a subset of samples \mathbf{T} (training data set) whose class vector is known, \mathbf{c}^{obs} , finding the subset of genetic signatures \mathbf{g} that maximizes the learning accuracy, that is, the number of samples whose class is correctly predicted. This can be written as the result of the following optimization problem:

$$O(\mathbf{g}) = \|\mathbf{L}^*(\mathbf{g}) - \mathbf{c}^{\text{obs}}\|_1, \quad (2)$$

$$\mathbf{L}^*(\mathbf{g}) = (\mathbf{L}^*(\mathbf{g}_1), \mathbf{L}^*(\mathbf{g}_2), \dots, \mathbf{L}^*(\mathbf{g}_l)), \quad (3)$$

where $\mathbf{L}^*(\mathbf{g})$ is the set of predicted classes, \mathbf{g}_i the genetic signature corresponding to the sample i in the training dataset \mathbf{T} , and $O(\mathbf{g})$ stands for the percentage of corrected predicted classes with respect to the total number of samples in \mathbf{T} . In practice, the predictive accuracy of a genetic signature (\mathbf{g}) is performed via cross-validation.

The generalization, that consists in predicting the class of a new sample (\mathbf{g}_{new}) whose class is unknown using the genetic signatures that have been found during the learning process. As noted above, phenotype prediction problems have a high degree of under-determinacy, since the number of monitored probes (or genes) is much greater than the number of samples (subjects enrolled for the clinical trial). Therefore, in the learning step, several lists of genes with similar predictive accuracy exist.

The set of high predictive networks is defined as $\mathbf{M}_{\text{tol}} = \{\mathbf{g} : O(\mathbf{g}) < \text{tol}\}$, and it is related to the uncertainty space of the classifier $\mathbf{L}^*(\mathbf{g})$, considering the optimization of the $O(\mathbf{g})$ as an inverse problem. For the same number of genes (given by s), these networks are located in one or several flat curvilinear valleys of the cost function topography [2, 3]. These signatures with different length (i.e. number of genes) are expected to be involved in the mechanistic pathways that explain the disease development. Within this set, the smallest-scale signature is the one that has the least number of discriminatory genes with the highest predictive accuracy. Such knowledge is important for early diagnosis, treatment and therapeutic target optimization.

Besides, due to noise in the genetic data and class assignment, it is possible that highly discriminatory signatures might be misinterpreted and not be directly involved in the genetic pathways [4]. Therefore, robust methods are needed to sample these genetic networks.

Bayesian Networks (BNs) have been used to identify genetic interactions in genetic experiments [5] and discover their relations with other kind of variables [6]. BNs are factored representations of probability distributions as stochastic directed acyclic graphs whose nodes and edges denote the existing relationships between variables. BNs follows a Bayesian approach of inverse problems, in which:

$$P(\mathbf{g}/\mathbf{c}^{\text{obs}}) = P(\mathbf{g})L(\mathbf{c}^{\text{obs}}/\mathbf{g})/P(\mathbf{c}^{\text{obs}}). \quad (4)$$

$P(\mathbf{g})$ is the prior distribution to sample the genetic signatures; $P(\mathbf{c}^{\text{obs}})$ is called the evidence of the observed class and it is considered as a normalization constant; $L(\mathbf{c}^{\text{obs}}/\mathbf{g})$ is called the likelihood of the genetic signature \mathbf{g} and depends on $O(\mathbf{g})$. That way, a genetic signature is more alike if it explains better the observed class, \mathbf{c}^{obs} , according to the prior sampling distribution. This way of reasoning does not take into account the role of the noise in the observed data (\mathbf{c}^{obs} and genetic data) that might falsify this hypothesis [4]. Finally, $P(\mathbf{g}/\mathbf{c}^{\text{obs}})$ is the posterior distribution of the genetic signatures, which is the main result of this procedure. BNs have the drawback of being very computationally demanding, mainly when the number of genes candidates to be sampled increases.

In this paper, we introduce the Fisher's ratio sampler that provides functional advantages in terms of computational costs and simplicity over a BNs approach, since the Fisher's ratio sampler does need to infer the posterior probability distribution of the most discriminatory genetic probes, and only samples the uncertainty space using a frequentist approach.

We show its application in identifying pathways involved in the metastasis of Triple Negative Breast Cancers (TNBC). TNBC refers to any breast cancer characterized by the absence of Estrogen Receptors (ER), Progesterone Receptors (PR) and Human Epidermal Growth factor 2 receptors (HER2). This exercise has clinical relevance as the absence of these three receptors significantly reduces targeted treatment options for patients with TNBC. TNBC accounts for somewhere between 10% and 17% of all breast cancers. The TNBC phenotype shows high metastatic potential and poor prognosis. The treatment for this type of cancer consists of adjuvant chemotherapy and radiotherapy. Unfortunately, response to chemotherapy does not correlate with overall survival. In addition, recurrences are observed in TNBC during the first and third years after therapy, and most deaths take place in the first five years. The survival decreases after the first metastatic event [7]. Therefore, in this heterogeneous group of tumors, new techniques and algorithms are necessary to better predict diagnosis and prognosis in order to establish appropriate therapies and improve patient survival. The understanding the molecular mechanisms that are involved in TNBC development is crucial to design new drugs.

For that purpose, we have modelled the microarray data deposited in the Gene Expression Omnibus (GEO) under the accession number GSE58812. This dataset

contains the microarray analysis of 107 patients diagnosed with TNBC and controlled for metastasis (44 relapsed and 63 were disease-free after a follow-up period of 7 years), that were studied and treated between 1998 and 2007 at the Institut de Cancérologie de l'Ouest – René Gauducheau and the Institut de Cancérologie de l'Ouest – Paul Papin. This data received the consent of patients as required by the French Committee for the Protection of Human Subjects (CCPPRB). Gene expression analysis were performed in quality control RNA samples using Affymetrix® Human Genome U133 Plus 2.0 Arrays (Affymetrix®, Santa Clara, CA, USA), and the microarray analysis measured over 43,000 transcripts [8]. In this case, the observed class c^{obs} consists in a binary flag indicating if the patient has or not developed metastasis, and the Fisher's ratio sampler aims at unravelling the altered genetic pathways in the metastasis events. This kind of classification problems have a very high underdetermined character, as pointed in [1].

2 The Fisher's Random Sampler

The Fisher's ratio sampler takes into account the discriminatory power of the differentially expressed genes as measured by the Fisher's ratio to induce the prior sampling distribution of the high discriminatory genetic networks in a phenotype prediction problem. The flowchart of the Fisher sampler algorithm is shown in Fig. 1.

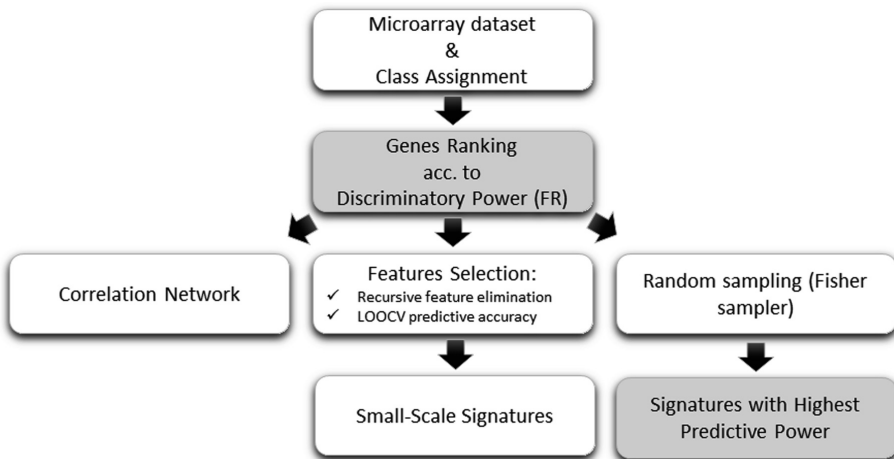


Fig. 1. Flow diagram of the Fisher's sampler algorithm.

The algorithm works as follows:

1. We first identified the set of genes with the highest Fisher's ratio from the set of genes with the highest fold change. For that purpose, we found the genes that were differentially expressed in both tails (over and under-expressed) and ranking these

genes by their Fisher's ratio to detect those that best separate the classes further apart and are very homogeneous within classes (low intra-class variance). In a binary classification problem the Fisher's ratio of the gene j is defined as:

$$FR_j = \frac{(\mu_{j1} - \mu_{j2})^2}{\sigma_{j1}^2 + \sigma_{j2}^2},$$

where μ_{ji} is a measure of the center of mass of the probability distribution of the gene j in class i , and σ_{ji} is a measure of its dispersion within this class. Discriminatory genes correspond to higher Fisher's ratios since they have a low intra-class dispersion (intra-class homogeneity) and high inter-class distance that accounts for the separation between the centers of the corresponding prognostic genes distributions.

We defined the set of discriminatory genes as those which were differentially expressed and had a Fisher's ratio greater than $f_r = 0.8$, since this value implies that the centers of the distribution in both classes are separated: $|\mu_{j1} - \mu_{j2}| > 0.89 \sqrt{\sigma_{j1}^2 + \sigma_{j2}^2}$. The Fisher's ratio cutoff value could be further decreased till $f_r = 0.5$ (or less) if the number of discriminatory genes within this set is very low. Therefore, the Fisher's ratio cut-off value is an important tuning parameter of this procedure.

2. Finding the small-scale genetic signature. The most discriminatory genes are ranked in decreasing order based on their Fisher's ratio and the algorithm finds the small-scale signature that optimally discriminates between classes by means of recursive feature elimination. The predictive accuracy estimation is based on Leave-One-Out-Cross-Validation (LOOCV) using a nearest neighbor classifier [1]. The small-scale signature provides an approximate idea of the typical length (number of genes) of the high discriminatory networks that are going to be sampled, and are related to the phenotype problem complexity.
3. Random sampling of high discriminatory equivalent networks. The random sampler is able to find other networks of highly discriminatory genes, using a prior sampling probability of any individual gene proportional to its Fisher's ratio. Genes are rated as *Headers* (variables with $FR > Tol$) and *Helpers* (variables with $FR < Tol$), as shown in Fig. 2, and the Fisher sampler builds genes signatures at each step, by selecting some Headers and some Helpers that meet the following conditions:

$$\begin{aligned} \max(cdf(Headers) < rand(1)) \\ \max(cdf(Helpers) < rand(1)); \end{aligned}$$

where cdf is the empirical cumulative distribution function in the sets of genes *Headers* and *Helpers* sets respectively, and $rand(1)$ is a random number between 0 and 1.

This way of proceeding is based on the fact that high discriminatory variables serve to span the main features of the classification, whilst variables with lowest discriminatory ratios account for the details in the discrimination. This method determines the minimum amount of high-frequency details (helper genes) that are

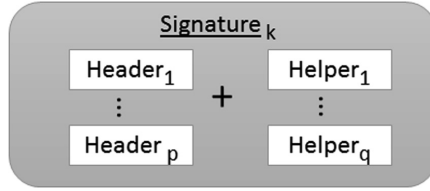


Fig. 2. Signature of prognostic genes built by the Fisher sampler at step k .

needed to optimally discriminate between classes promoting the header genes, which are those that explain the phenotype in a robust manner.

Once a network is randomly built according to the Fisher’s probability distribution, its LOOCV predictive accuracy is established. This sampling follows the Bayes’ rule with a prior probability that depends on the Fisher’s ratio of the genes that have been selected, and using a likelihood $L(\mathbf{c}^{\text{obs}}/\mathbf{g})$ that depends on the LOOCV predictive accuracy of this network. Nevertheless, in this case we are interested in factoring the posterior distribution, but in unravelling the sets of genes with highest discriminatory power that are related to the uncertainty space of the phenotype prediction problem [1–3] via $\mathbf{L}^*(\mathbf{g})$, and identifying the altered genetic pathways via these sets.

4. Finally, taking into account the most discriminatory networks that have been sampled, determining those that best defined the TNBC phenotype with a predictive accuracy higher than a minimum accuracy that it is given (typically higher than 90%). If needed this threshold could be decrease to 85%. Based on these networks we establish the posterior sampling frequencies of the main prognostic genes. The biological pathways are identified using Gene Analytics [9] by querying the group of genes with the higher sampling frequency. The frequency threshold was selected to optimize the discriminatory genes in the pathway analysis. This platform also provides important clues about the biological processes and the existence of chemical compounds to target the actionable genes.

A simplified version of this algorithm has been previously used to predict the genomic risk of aromatase inhibitor-related arthralgia in patients with breast cancer using SNPs [10], and also in the integration of genomic data to analyze the stability of the genetic signatures that predict different mutations in CLL patients [11].

3 Results

Table 1 shows the most-frequently sampled genes by the Fisher’s ratio sampler. All these genes are overexpressed in the metastasis group. The frequency is calculated over the total number of genes sampled in the high predictive networks with accuracy higher than 90%, found after 10000 random simulations.

This algorithm is very fast, since these simulations took less than 5 min in a personal computer. The most important gene is LINC00630 which is overexpressed in the metastasis group. Mao et al. [12] studied the role of non-coding RNAs showing that

Table 1. List of most-frequently sampled genes by the FR sampler (means in log₂ scale).

Frequency (%)	Gene name	Mean-metastasis	Mean-no metastasis
0.75	LINC00630	15	8
0.47	LOC100506272	23	16
0.45	STC1	382	263
0.42	BAIAP2-AS1	24	16
0.39	ARFGAP2	31	22
0.38	LHX9	19	15
0.36	LOC646482	32	24
0.36	CACNA1S	57	38
0.36	AC108056.1	17	13
0.36	NXF3	23	15
0.34	GIPC3	60	38
0.34	KCNS2	27	16
0.34	DAZ1	16	10
0.34	UGT1A1	39	26
0.34	RP5-855D21.1	30	19
0.33	EXOC5	15	10
0.33	MDM2	28	20
0.33	CXADR	5891	4070
0.33	RP3-333B15.4	66	48

LINC00630 play a critical role in the development of Non-Small-Cell Lung Cancers, showing that Linc00630 overexpression increased cell proliferation and metastasis in vitro and in vivo whereas LINC00630 silencing had opposite effects. The gene STC1 (Stanniocalcin-1) encodes a glycoprotein that is expressed in a wide variety of tissues. Overexpression of STC1 in mice produces high serum phosphate levels, dwarfism, and increased metabolic rate. This gene has altered expression in hepatocellular, ovarian, and breast cancers. In TNBC this gene has been associated to metastasis [13]. BAIAP2-AS1 is a non-coding RNA gene. This gene has been associated to hepatitis B virus-related hepatocellular carcinoma [14]. Metabolism related items and cancer associated KEGG pathways are in relation with BAIAP2-AS. It has been found that BAIAP2-AS1 may function as a ceRNA. ARFGAP2 (ADP Ribosylation Factor GTPase Activating Protein 2) is a Protein coding gene that it is involved in protein recycling (Transport of the damaged proteins to the Golgi and subsequent modification). LHX9 is a gene involved in transcription. The most important pathways identified by Gene Analytics involve P53 and RhoA Signaling and DREAM Repression and Dynorphin Expression. Although the individual ontological attributions of the most important genes in the prediction of the TNBC phenotype is very interesting and shows the important pedigree of these genes, the most important feature is to understand how the work in synergy. Table 2 shows the most important pathways and biological processes involved, which are: Phagocytosis, Positive Regulation of B Cell Activation and Extrinsic Apoptotic Signaling Pathway.

Table 2. List of most important pathways and biological processes found by the FR sampler.

Score	Top pathways
10.3	Direct P53 effectors
10.1	DREAM repression and dynorphin expression
9.6	P53 signaling
8.8	RhoA signaling pathway
8.4	P53 pathway (RnD)
Score	Top biological processes
11.4	Phagocytosis, recognition
11.4	Positive regulation of B cell activation
11.2	Extrinsic apoptotic signaling pathway
9.7	Phagocytosis, engulfment
9.5	B cell receptor signaling pathway

Interestingly, the top biological process involved (phagocytosis) is a major mechanism in the immune system defense which is related to NF- κ B activation.

4 Conclusions

In this paper, we have introduced the Fisher's ratio sampler to analyze the defective pathways in phenotype prediction problems that are highly underdetermined. This novel sampling algorithm is able to sample the defective pathways using a prior probability distribution that is proportional to the Fisher's ratio of the differentially expressed genes. These genes are the headers in the phenotype discrimination and should be considered as main therapeutically targets.

This algorithm is very robust and it is not computationally expensive compared to BN. We show the application to the analysis of the transcriptome in a cohort of women diagnosed with Triple Negative Breast Cancers in France, showing the importance of P53 and RhoA Signaling pathways and Phagocytosis biological processes. This sampling methodology could be also applied to other diseases and any kind of genetic data (microarrays, RNA-seq, SNPs, etc.). We have chosen the TNBC phenotype due to its high metastatic potential and poor prognosis. We hope that the analysis shown in this paper will provide further insight and new therapeutic targets. The Fisher's ratio sampler is an agnostic methodology that can be also applied to the genetic analysis of other complex diseases.

References

1. De Andrés Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Design of biomedical robots for phenotype prediction problems. *J. Comput. Biol.* **23**(8), 678–692 (2016)
2. Fernández-Martínez, J.L., Fernández-Muñiz, M.Z., Tompkins, M.J.: On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics* **77**(1), W1–W15 (2012). <https://doi.org/10.1190/geo2011-0341.1>

3. Fernández-Martínez, J.L., Pallero, J.L.G., Fernández-Muñiz, Z., Pedruelo-González, L.M.: From Bayes to Tarantola: new insights to understand uncertainty in inverse problems. *J. Appl. Geophys.* **98**, 62–72 (2013)
4. De Andrés-Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Sensitivity analysis of gene ranking methods in phenotype prediction. *J. Biomed. Inf.* **64**, 255–264 (2016)
5. Jiang, X., Barmada, M.M., Visweswaran, S.: Identifying genetic interactions in genome-wide data using Bayesian networks. *Genet. Epidemiol.* **34**(6), 575–581 (2010)
6. Su, C., Andrew, A., Karagas, M.R., Borsuk, M.E.: Using Bayesian networks to discover relations between genes, environment, and disease. *BioData Mining* **6**, 6 (2013)
7. Liedtke, C., Mazouni, C., et al.: Response to neoadjuvant therapy and long-term survival in patients with triple negative breast cancer. *J. Clin. Oncol.* **26**(8), 1275–1281 (2008)
8. Jézéquel, P., Loussouarn, D., Guérin-Charbonnel, C., Champion, L., et al.: Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* **17**, 43 (2015)
9. Stelzer, G., Inger, A., Olender, T., Iny-Stein, T., Dalah, I., Harel, A., et al.: GeneDecks: paralog hunting and gene-set distillation with GeneCards annotation. *OMICS* **13**(6), 477 (2009)
10. Reinbolt, R.E., Sonis, S., Timmers, C.D., Fernández-Martínez, J.L., Cernea, A., de Andrés-Galiana, E.J., Hashemi, S., Miller, K., Pilarski, R., Lustberg, M.B.: Genomic risk prediction of aromatase inhibitor-related arthralgia in patients with breast cancer using a novel machine-learning algorithm. *Cancer Med.* (2017). <https://doi.org/10.1002/cam4.1256>
11. Fernández-Martínez, J.L., de Andrés-Galiana, E.J., Sonis, S.T.: Genomic data integration in chronic lymphocytic leukemia. *J. Gene Med.* **19** (2017). <https://doi.org/10.1002/jgm.2936>
12. Mao, G., Jin, H., Wu, L.: DDX23-Linc00630-HDAC1 axis activates the Notch pathway to promote metastasis. *Oncotarget* **8**(24), 38937–38949 (2017)
13. Jeon, M., Han, J., Nam, S.J., Lee, J.E., Kim, S.: STC-1 expression is upregulated through an Akt/NF- κ B-dependent pathway in triple-negative breast cancer cells. *Oncol. Rep.* **36**(3), 1717–1722 (2016)
14. Gong, X., Wei, W., Chen, L., Xia, Z., Yu, C.: Comprehensive analysis of long non-coding RNA expression profiles in hepatitis B virus-related hepatocellular carcinoma. *Oncotarget* **7** (27), 42422–42430 (2016)



Sampling Defective Pathways in Phenotype Prediction Problems via the Holdout Sampler

Juan Luis Fernández-Martínez^{1(✉)}, Ana Cernea¹,
Enrique J. deAndrés-Galiana^{1,2}, Francisco Javier Fernández-Ovies¹,
Zulima Fernández-Muñiz¹, Oscar Alvarez-Machancoses¹,
Leorey Saligan³, and Stephen T. Sonis^{4,5}

¹ Group of Inverse Problems, Optimization and Machine Learning,
Department of Mathematics, University of Oviedo,
C/Federico García-Lorca, 18, 33007 Oviedo, Spain
jlfm@uniovi.es

² Department of Informatics and Computer Science,
University of Oviedo, Oviedo, Spain

³ National Institutes of Health, National Institute of Nursing Research,
Bethesda, MD, USA

⁴ Primary Endpoint Solutions, Watertown, MA, USA

⁵ Brigham and Womens' Hospital and the Dana-Farber Cancer Institute,
Boston, MA, USA

Abstract. In this paper, we introduce the holdout sampler to find the defective pathways in high underdetermined phenotype prediction problems. This sampling algorithm is inspired by the bootstrapping procedure used in regression analysis to established confidence bounds. We show that working with partial information (data bags) serves to sample the linear uncertainty region in a simple regression problem, mainly along the axis of greatest uncertainty that corresponds to the smallest singular value of the system matrix. This procedure applied to a phenotype prediction problem, considered as a generalized prediction problem between the set of genetic signatures and the set of classes in which the phenotype is divided, serves to unravel the ensemble of altered pathways in the transcriptome that are involved in the disease development. The algorithm looks for the minimum-scale genetic signature in each random holdout and the likelihood (predictive accuracy) is established using the validation dataset via a nearest-neighbor classifier. The posterior analysis serves to identify the header genes that most-frequently appear in the different hold-outs and are therefore robust to a partial lack of samples. These genes are used to establish the genetic pathways and the biological processes involved in the disease progression. This algorithm is much faster, robust and simpler than Bayesian Networks. We show its application to a microarray dataset concerning a type of breast cancers with poor prognoses (TNBC).

1 Introduction

The accuracy of genomic predictions of phenotype are often limited by the computational limitations associated with the excessive ratio of monitored genetic probes to the number of collected samples (patients). This fact, which is especially noteworthy as the number of

probes per microarray continues to grow, provokes a high dimensional uncertainty space resulting in ambiguity in the identification and characterization of biological pathways associated with the phenotype [1]. The impact of this ambiguity is potentially significant as it might provoke the adoption of wrong conclusions that could be associated with drug targeting, phenotype risk and patient responsiveness to treatment.

Therefore, novel robust methods are needed to characterize the altered genes and the genetic pathways in which they participate and that are involved in the disease development. In this paper, we present the holdout sampler that it is based in the method of bootstrapping or random sampling with replacement [2]. This statistical technique is used to establish the confidence intervals in sample estimates and, also to estimate the sampling distribution of any statistic using random sampling methods.

Let us suppose that we have built a classifier $\mathbf{L}^*(\mathbf{g})$ that relates the different genetic signatures to the class prediction in which the phenotype is divided:

$$\mathbf{L}^*(\mathbf{g}) : \mathbf{g} \in \mathbb{R}^s \rightarrow \mathbf{C} = \{c_1, c_2\}. \quad (1)$$

Finding the discriminatory genetic signatures corresponding to $\mathbf{L}^*(\mathbf{g})$ can be interpreted as a generalized regression model of the observed class vector \mathbf{c}^{obs} in the training data set \mathbf{T} with respect to the genetic signatures, in order to maximize the learning accuracy, that is, the percentage of corrected predicted classes via cross-validation:

$$\mathbf{O}(\mathbf{g}) = \|\mathbf{L}^*(\mathbf{g}) - \mathbf{c}^{\text{obs}}\|_1, \quad (2)$$

$$\mathbf{L}^*(\mathbf{g}) = (\mathbf{L}^*(\mathbf{g}_1), \mathbf{L}^*(\mathbf{g}_2), \dots, \mathbf{L}^*(\mathbf{g}_1)), \quad (3)$$

where $\mathbf{L}^*(\mathbf{g})$ is the set of predicted classes of the samples in the training dataset \mathbf{T} . The term $\|\mathbf{L}^*(\mathbf{g}) - \mathbf{c}^{\text{obs}}\|_1$ in Eq. (2) is the L_1 norm of the observed classes mismatch, that coincides in this cases with number of incorrect predicted samples.

Any phenotype prediction problem always has a high degree of under-determinacy, since the number of monitored probes (or genes) is much greater than the number of samples (subjects enrolled for the clinical trial). Therefore, several lists of genes with similar predictive accuracy exist. The set of high predictive networks is defined as $\mathbf{M}_{\text{tol}} = \{\mathbf{g} : \mathbf{O}(\mathbf{g}) < \text{tol}\}$, and it is related to the uncertainty space of $\mathbf{L}^*(\mathbf{g})$ in the phenotype prediction problem. For a given number of genes (or genetic length) these networks are located in several flat curvilinear valleys of the cost function topography [3, 4]. We will illustrate numerically with a simple regression problem that the simplest way of sampling its uncertainty region is via the holdout sampler described in this paper.

2 From Regression Analysis to Sampling the Genetic Pathways

Let us consider a simple linear regression problem of the kind $y^*(x) = a_0 + a_1x$, to a dataset given by $\{(x_1, y_1), (x_2, y_2), \dots, (x_s, y_s)\}$. The least squares problem consists of finding the parameters $\mathbf{m} = (a_0, a_1)$ so that the distance between the observed data

$$\mathbf{y}^{\text{obs}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_s \end{pmatrix} \text{ and the corresponding predictions } \mathbf{y}^{\text{pre}}(\mathbf{m}) = \begin{pmatrix} y_1^*(\mathbf{m}) \\ y_2^*(\mathbf{m}) \\ \vdots \\ y_s^*(\mathbf{m}) \end{pmatrix} \text{ is minimum}$$

according to the Euclidean distance in \mathbb{R}^s . The uncertainty analysis consists in sampling the family of equivalent models \mathbf{m} that fit the observed data \mathbf{y}^{obs} within the same error bounds:

$$\mathbf{M}_{\text{tol}}(\mathbf{m}) = \left\{ \mathbf{m} : E(\mathbf{m}) = \|\mathbf{y}^{\text{pre}}(\mathbf{m}) - \mathbf{y}^{\text{obs}}\|_2 < \text{tol} \right\}. \quad (4)$$

Typically, the error function is expressed in relative terms to speak about %. Fernández-Martínez et al. [3] proved that the topography of the data error cost function corresponds to a straight flat elongated valley if the inverse problem is linear, whereas in the nonlinear case, the cost function topography consists of one or more curvilinear valleys (or basins) of low misfits eventually connected by saddle points [3, 4].

Besides, in the linear regression case the equivalent models belong to an ellipse whose axes and orientations are related to the ill-conditioning of the system matrix $\mathbf{F} = [\mathbf{1}_{\mathbb{R}^s}, \mathbf{x}]$. The least-squares model solves the linear system of normal equations, $\mathbf{F}^T \mathbf{F} \mathbf{m} = \mathbf{F}^T \mathbf{y}^{\text{obs}}$, and provides the model parameters \mathbf{m} with minimum prediction error (data misfit). The region of uncertainty of value *tol* is an ellipse whose axes are oriented following the eigenvectors of the matrix $\mathbf{F}^T \mathbf{F}$ with lengths proportional to the inverse of the square roots of the eigenvalues of this matrix. That way, the uncertainty is higher in the direction with the smallest eigenvalue. Besides, when matrix \mathbf{F} has a non-trivial kernel, then the ellipse degenerates in two parallel lines and the axis of the uncertainty region in the direction of the kernel of \mathbf{F} is infinite.

A simple way of sampling the posterior distribution of the model parameters in this linear regression problem consists in performing least-squares of different data bags. It is important to remark that in this simple linear regression case, all the least-squares systems that are solved are overdetermined. Therefore, no overfitting of the observed data is done. For instance, Fig. 1 shows for a simple linear regression model the ellipse of uncertainty for a relative misfit lower than 15% and the different sets of parameters $\{\mathbf{a}_{0k}, \mathbf{a}_{1k}\}_{k=1, \dots, N_s}$ obtained in the least squares fitting of different N_s bagging experiments, where the training and validation sets are randomly generated. It can be observed that these model parameters are aligned along the axis of maximum uncertainty. Therefore, this bagging (or bootstrapping) procedure serves to sample the uncertainty region of this simple linear regression problem, since the use of partial information in the misfit procedure induces uncertainty in the model parameter identification. Obviously, we are only interested in the model parameters that fit the validation set within the same error bounds given by the *tol* parameter. These models are called equivalent, since they explain the observed data within similar error bounds.

This simple example provides the basis for the hold-out sampler, but the main difference with respect to phenotype prediction problems is the high degree of underdeterminacy of the latter, due to their large number of genes that are monitored and could be potential targets. Therefore, some additional ranking methods should be

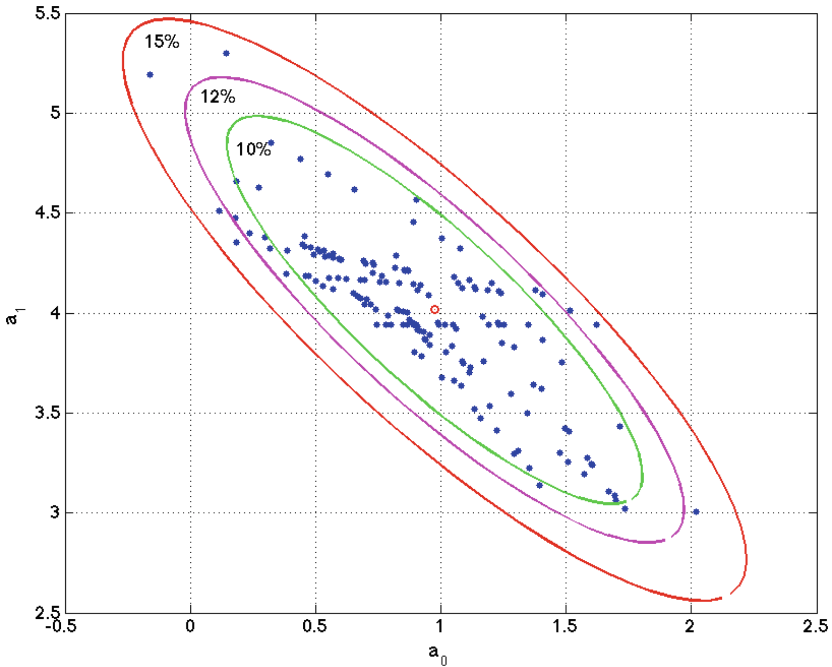


Fig. 1. Linear regression model. Ellipse of uncertainty for a relative misfit of 15% and different sets of model parameters found in the different bagging experiment. It can be observed that these models sample the region of uncertainty within the ellipse of 15% relative misfit. This example is very important to understand that the list of genes that equally explain a phenotype is not unique, and one simple method to sample these high discriminatory genetic networks is by performing random holdout, looking for the minimum-scale signatures that better explain the phenotype in each hold out, and finding the most-frequently sampled genes in these signatures, that are similar in phenotype prediction problems to the points (model parameters) located within the ellipse of this simple regression problem.

applied to reduce the dimension of the model space (genetic probes) and avoiding noise overfitting. Besides, in the phenotype prediction problem no least-squares system is solved, since the method to predict the observed classes is based in gene selection via a k-NN classifier. More details are given below.

The implications of the previous result in the analysis of defective (or deregulated/altered) pathways in phenotype prediction problems is straight-forward since this bootstrapping methodology can be used to sample the header genes that are discriminatory in the different hold outs and performing posterior analysis of these networks to unravel the biological pathways that are involved in the disease development. These networks, that have different number of genes, are expected to be involved in the genetic pathways that explain the disease development. Within this set, the smallest-scale signature is the one that has the least number of discriminatory genes with the highest predictive accuracy.

This knowledge is very important for early diagnosis and treatment optimization. The main reason to accomplish the sampling of these networks is understanding the main causes in the disease development and being able to find new therapeutically targets. Due to noise in the genetic data and also to wrong class assignment of the samples by the medical experts (misdiagnosis), some of these highly discriminatory signatures might be falsified and not be directly involved in the genetic pathways [5].

Different methodologies have been proposed to rank the genes according to their discriminatory power [6]. In this paper, we use a combination of fold-change and Fisher's ratio due to its robustness [5]. Bayesian Networks (BNs) have been used to identify genetic interactions in genetic experiments [7], and subsequently define both pathways and ontology. The Bayesian approach of the phenotype prediction problems writes:

$$P(\mathbf{g}/\mathbf{c}^{\text{obs}}) \sim P(\mathbf{g})L(\mathbf{c}^{\text{obs}}/\mathbf{g}), \quad (5)$$

where $P(\mathbf{g})$ is the prior distribution to sample the genetic signatures; $L(\mathbf{c}^{\text{obs}}/\mathbf{g})$ is called the likelihood of \mathbf{g} and depends on $\mathbf{O}(\mathbf{g})$; and $P(\mathbf{g}/\mathbf{c}^{\text{obs}})$ is the posterior distribution of the genetic signatures, which is the main result of this approach. Bayesian Networks (BN) have the drawback of being very computationally demanding due to the large number of possible combinations (genetic probes).

In this paper, we present the hold-out sampler that is much faster and theoretically simple than BNs, since our approach is frequentist and uses a gene selection that serves to reduce the under determinacy of the phenotype prediction. Besides we are not interested in factorizing $P(\mathbf{g}/\mathbf{c}^{\text{obs}})$, just to sample the main genetic networks that constitute the uncertainty space of the phenotype prediction problem, hoping that this knowledge will help to better understand the mechanisms that are involved in the disease. The hold-out algorithm is very fast and can be run in a personal computer in real time. In the present case, we have performed 500 simulations in less than 5 min.

We show the application to the sample the pathways involved in the metastasis of Triple Negative Breast Cancers (TNBC). The TNBC phenotype shows high metastatic potential and in general a poor prognosis. For that purpose, we have modelled the microarray data deposited in the Gene Expression Omnibus (GEO) under the accession number GSE58812, which contains the microarray analysis of 107 French patients (44 relapsed and 63 were disease-free after a follow-up period of 7 years) diagnosed and treated between 1998 and 2007 [8].

The flowchart for the holdout sampler is as follows:

1. Data bagging: Different random 75/25 data bag holdouts were created from the dataset, where 75% of the data is used for learning and 25% for validation. In the present case, we have generated 1000 different bags. For each of these bags the genes were ranked according to their discriminatory power measured by the Fisher's ratio, maximizing the distance between the centers of the gene expression in each class, and minimizing at the same time the intra-class variance. The Fisher's ratio was combined with a previous fold-change analysis to avoid the genes with high FR and low fold-change due to small dispersions within each class. We define the set of discriminatory genes as those that are differentially expressed and have Fisher's ratio greater

than a minimum cut-off, set up to establish how the classes separate. The most discriminatory variables were ranked in decreasing order by their discriminatory power and with application of the algorithm to identify shortest list of prognostic variables with the highest predictive accuracy using recursive feature elimination (minimum-scale genetic signature). This way of proceeding is based on the fact that high discriminatory variables serve to span the main features of the classification, whilst variables with lowest discriminatory ratios account for the details in the discrimination. This method determines the minimum amount of high-frequency details (helper genes) that are needed to optimally discriminate between classes promoting the header genes, which are those that explain the phenotype in a robust manner. The predictive accuracy estimation is based on Leave-One-Out-Cross-Validation (LOOCV) by averaging the LOOCV predictive accuracy over all the samples of the validation dataset in each bag. This methodology involves a k-Nearest-Neighbor classifier in the reduced set of high discriminatory genes (minimum-scale signature) and has been successfully applied to the bioinformatics modeling of high dimensional Omics data [9, 10].

2. Posterior analysis: once the bags simulation has been finished, the posterior analysis consists in finding all the minimum size signatures that have a validation predictive accuracy higher than a given tolerance. In this case we have considered all the holdouts with predictive accuracy higher than 90%. This threshold is fixed high enough in order to have explicative genetic networks of the phenotype. This threshold could be decreased till 80–85% if the number of high predictive networks is too small (tens). Finally, we perform frequency analysis of these lists to find the most frequently sampled genes that serve to establish the defective genetic pathways using GeneAnalytics [11], which are those related to the deregulated genes. This software uses the main ontological databases and provides important clues about the biological processes involved.

3 Results

Table 1 shows the most-frequently sampled genes by the Fisher’s ratio sampler. Most of all these genes, but LOC644135, are overexpressed in the metastasis group. The most important gene appears to be STC1 (Stanniocalcin-1) with 2 different probes in the set of most important sampled genes. This gene encodes a glycoprotein and its overexpression produced high serum phosphate levels and the increase of the metabolic rate in mice. This gene has also an altered expression in hepatocellular, ovarian and breast cancers, and it has been previously associated to metastasis in TNBC [12, 13].

Other important gene is OTUB2 that encodes one enzyme, which is needed to reverse the ubiquitin modification of damaged proteins. Among its related pathways are ovarian tumor domain proteases and protein ubiquitination. BAIAP2-AS1 is a non-coding RNA gene, that been associated to hepatitis B virus-related hepatocellular carcinoma [14]. KCNS2 encodes a protein that is a voltage-gated potassium channel subunit. The importance of potassium channels in cancer has been reviewed by Huang and Jan [15], since they regulate cancer cell behaviors such as proliferation and migration. It has been shown that potassium channel activation inhibits proliferation of breast cancer cells [16].

Table 1. List of most-frequently sampled genes by the hold-out sampler (means are given in log₂ scale).

Frequency (%)	Gene name	Mean-metastasis	Mean-no metastasis
0.21	OTUB2	45	28
0.21	STC1	382	262
0.21	BAIAP2-AS1	24	16
0.21	STC1	1182	723
0.21	KCNS2	27	16
0.21	LOC100506272	23	16
0.21	LOC644135	14	26
0.21	LINC00630	15	8
0.2	UGT1A1	39	26
0.2	ARFGAP2	31	22
0.2	CACNA1I	57	38
0.2	DCAF8	40	26
0.19	RP11-799D4.4	15	25
0.19	MDM2	28	20
0.19	RP11-38C18.3	30	21
0.19	BFSP2-AS1	24	16

LOC100506272 and LOC644135 are two uncharacterized genes. LINC00630 is a non-coding gene. LINC00630 plays a critical role in the development of Non-Small-Cell Lung Cancers [17]. Overexpression of LINC00630 increases cell proliferation and metastasis, whereas LINC00630 silencing has opposite effects.

The most important pathways identified by GeneAnalytics are several signaling pathways, such as, JNK Signaling, RhoA Signaling, FoxO Signaling and TGF-beta Signaling Pathway. The most important biological processes involved are defense response, DNA Damage Response, response to Hypoxia and Response to Steroid Hormone (Table 2). These results are similar to those found by the Fisher's ratio sampler [18].

Table 2. List of most important pathways and biological processes found by the holdout sampler.

Score	Top pathways
11.16	JNK signaling in the CD4 + TCR pathway
9.71	RhoA signaling pathway
8.28	ATM pathway
8.06	FoxO signaling pathway
8.01	TGF-beta signaling pathway
Score	Top biological processes
12.22	Defense response
10.50	DNA Damage response, signal transduction by P53 class mediator resulting in cell cycle arrest
10.41	Response to hypoxia
10.40	Response to steroid hormone

4 Conclusions

In this paper, we have introduced the Holdout sampler to analyze the defective pathways in phenotype prediction problems that have a very highly underdetermined. This novel algorithm is able to sample the defective pathways by searching the minimum-scale signature of different random holdouts.

We show that this bootstrapping strategy serves to sample the uncertainty region of a simple linear regression problem. This algorithm is very robust and it is not computationally expensive compared to Bayesian Networks. We show the application to the analysis of the transcriptome in a cohort of women diagnosed with Triple Negative Breast Cancers in France, due to the high metastatic potential and poor prognosis of the TNBC. This analysis highlights the importance of several signaling pathways (JNK, RhoA, FoxO and TGF-beta), and defense response, DNA damage, response to hypoxia and to steroid hormone as the main biological processes involved. This sampling methodology could be also applied to other diseases and different types of genetic data (microarrays, RNA-seq, SNPs, etc.). We think that its application to complex genetic datasets opens new possibilities in translational medicine and drug design.



References

1. de Andrés Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Design of biomedical robots for phenotype prediction problems. *J. Comput. Biol.* **23**(8), 678–92 (2016)
2. Efron, B., Tibshirani, R.: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, Boca Raton (1993). ISBN 0-412-04231-2
3. Fernández-Martínez, J.L., Fernández-Muñiz, M.Z., Tompkins, M.J.: On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics* **77**(1), W1–W15 (2012). <https://doi.org/10.1190/geo2011-0341.1>
4. Fernández-Martínez, J.L., Pallero, J.L.G., Fernández-Muñiz, Z., Pedruelo-González, L.M.: From Bayes to Tarantola: new insights to understand uncertainty in inverse problems. *J. Appl. Geophys.* **98**, 62–72 (2013)
5. de Andrés-Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Sensitivity analysis of gene ranking methods in phenotype prediction. *J. Biomed. Inform.* **64**, 255–264 (2016)
6. Saeys, Y., Inza, I., Larrañaga, P.: A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**(19), 2507–2517 (2007)
7. Jiang, X., Barmada, M.M., Visweswaran, S.: Identifying genetic interactions in genome-wide data using Bayesian networks. *Genet. Epidemiol.* **34**(6), 575–581 (2010)
8. Jézéquel, P., Loussouarn, D., Guérin-Charbonnel, C., Campion, L., et al.: Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* **20**(17), 43 (2015)
9. Saligan, L.N., Fernández-Martínez, J.L., de Andrés Galiana, E.J., Sonis, S.: Supervised classification by filter methods and recursive feature elimination predicts risk of radiotherapy-related fatigue in patients with prostate cancer. *Cancer Inform.* **13**(141–152), 2014 (2014)
10. Fernández-Martínez, J.L., de Andrés-Galiana, E.J., Sonis, S.: Genomic data integration in chronic lymphocytic leukemia. *J. Gene Med.* **19**, 1–2 (2017)

11. Stelzer, G., Inger, A., Olender, T., Iny-Stein, T., Dalah, I., Harel, A., et al.: GeneDecks: paralog hunting and gene-set distillation with GeneCards annotation. *OMICS* **13**(6), 477 (2009)
12. Jeon, M., Han, J., Nam, S.J., Lee, J.E., Kim, S.: STC-1 expression is upregulated through an Akt/NF- κ B-dependent pathway in triple-negative breast cancer cells. *Oncol. Rep.* **36**(3), 1717–1722 (2016). Epub 25 July 2016
13. Han, J., Jeon, M., Shin, I., Kim, S.: Elevated STC-1 augments the invasiveness of triple-negative breast cancer cells through activation of the JNK/c-Jun signaling pathway. *Oncol. Rep.* **36**(3), 1764–71 (2016). Epub 26 July 2016
14. Gong, X., Wei, W., Chen, L., Xia, Z., Yu, C.: Comprehensive analysis of long non-coding RNA expression profiles in hepatitis B virus-related hepatocellular carcinoma. *Oncotarget* **7** (27), 42422–42430 (2016). <http://doi.org/10.18632/oncotarget.9880>
15. Huang, X., Jan, L.Y.: Targeting potassium channels in cancer. *J. Cell Biol.* **206**(2), 151–162 (2016). <https://doi.org/10.1083/jcb.201404136>
16. Lansu, K., Gentile, S.: Potassium channel activation inhibits proliferation of breast cancer cells by activating a senescence program. *Cell Death Dis.* **4**, e652 (2013). <https://doi.org/10.1038/cddis.2013.174>
17. Mao, G., Jin, H., Wu, L.: DDX23-Linc00630-HDAC1 axis activates the Notch pathway to promote metastasis. *Oncotarget.* **8**(24), 38937–38949 (2017). <https://doi.org/10.18632/oncotarget.17156>
18. Cernea, A., Fernández-Martínez, J.L., de Andrés-Galiana, E.J., Fernández-Ovies, F.J., Fernández-Muñiz, Z., Álvarez-Machancoses, O., Saligan, L., Sonis, S.: Sampling defective pathways in phenotype prediction problems via the Fisher’s ratio sampler. In: *IWBBIO 2018* (2018)



Comparison of Different Sampling Algorithms for Phenotype Prediction

Ana Cernea¹ , Juan Luis Fernández-Martínez¹ ,
Enrique J. de Andrés-Galiana^{1,2}, Francisco Javier Fernández-Ovies¹,
Zulima Fernández-Muñiz¹, Óscar Alvarez-Machancoses¹,
Leorey Saligan³, and Stephen T. Sonis^{4,5}

¹ Group of Inverse Problems, Optimization and Machine Learning,
Department of Mathematics, University of Oviedo, C/Federico García-Lorca,
18, 33007 Oviedo, Spain

jlfm@uniovi.es

² Department of Informatics and Computer Science, University of Oviedo,
Oviedo, Spain

³ National Institute of Nursing Research, National Institutes of Health,
Bethesda, MD, USA

⁴ Primary Endpoint Solutions, Watertown, MA, USA

⁵ Brigham and Womens¹ Hospital and the Dana-Farber Cancer Institute,
Boston, MA, USA

Abstract. In this paper, we compare different sampling algorithms used for identifying the defective pathways in highly underdetermined phenotype prediction problems. The first algorithm (Fisher's ratio sampler) selects the most discriminatory genes and samples the high discriminatory genetic networks according to a prior probability that it is proportional to their individual Fisher's ratio. The second one (holdout sampler) is inspired by the bootstrapping procedure used in regression analysis and uses the minimum-scale signatures found in different random hold outs to establish the most frequently sampled genes. The third one is a pure random sampler which randomly builds networks of differentially expressed genes. In all these algorithms, the likelihood of the different networks is established via leave one out cross-validation (LOOCV), and the posterior analysis of the most frequently sampled genes serves to establish the altered biological pathways. These algorithms are compared to the results obtained via Bayesian Networks (BNs). We show the application of these algorithms to a microarray dataset concerning Triple Negative Breast Cancers. This comparison shows that the Random, Fisher's ratio and Holdout samplers are most effective than BNs, and all provide similar insights about the genetic mechanisms that are involved in this disease. Therefore, it can be concluded that all these samplers are good alternatives to Bayesian Networks which much lower computational demands. Besides this analysis confirms the insight that the altered pathways should be independent of the sampling methodology and the classifier that is used to infer them.

1 Introduction

Phenotype prediction is still a main challenge in biomedical research and consists of identifying the set(s) of genes that influence the disease genesis and development. However, this kind of prediction problems are highly underdetermined since the number of monitored genetic probes is always much larger than the number of observed samples. Consequently, the associated uncertainty space of these problems is huge, and the characterization of the involved biological pathways is very ambiguous because there exist many equivalent genetic networks that predict the phenotype with similar accuracies [1, 2].

As well, one of the main obstacles in the analysis of genetic data is the absence of a conceptual model that relates the different genes/probes to the class prediction, between the set of genetic signatures \mathbf{g} and the set of classes $C = \{1, 2\}$ in which the phenotype is divided:

$$\mathbf{L}^*(\mathbf{g}): \mathbf{g} \in \mathbb{R}^s \rightarrow C = \{1, 2\}. \quad (1)$$

Finding the discriminatory genetic signatures corresponding to $\mathbf{L}^*(\mathbf{g})$ implies the optimization of the the cost function $O(\mathbf{g}) = \|\mathbf{L}^*(\mathbf{g}) - \mathbf{c}^{\text{obs}}\|_1$, that involves the observed classes (\mathbf{c}^{obs}) of a set of samples in the training dataset, and the corresponding set of predictions $\mathbf{L}^*(\mathbf{g})$, using the classifier \mathbf{L}^* , via the genetic signature \mathbf{g} that has to found. $\|\mathbf{L}^*(\mathbf{g}) - \mathbf{c}^{\text{obs}}\|_1$ represents the L_1 norm of the prediction error, which coincides in this cases with the number of uncorrected sampled predicted by the classifier. The uncertainty space relative to \mathbf{L}^* , $M_{\text{tol}} = \{\mathbf{g}: O(\mathbf{g}) < \text{tol}\}$, is composed the sets of high predictive networks with similar predictive accuracy. These sets are located in one or several flat curvilinear valleys of the cost function topography, $O(\mathbf{g})$ [3]. The main reason to accomplish the sampling of these networks is to understand the causes of disease development and finding new therapeutic targets. This knowledge concerns (in principle) the use of the classifier $\mathbf{L}^*(\mathbf{g})$.

In this paper, we compare different sampling methods that involve different classifiers to establish those that seem to be more robust in the genetic pathway identification.

The first algorithm concerns the Fisher's ratio (FR) sampler [4], which consists in sampling the defective pathways taking into account the discriminatory power of the differentially expressed genes measured by their Fisher's ratio, that it is used to infer a prior sampling distribution of the high discriminatory networks. The sampled networks are randomly established using this prior distribution, and their likelihood (or predictive accuracy) is established via LOOCV using a nearest-neighbor classifier [5]. This sampling algorithm, that implies a frequentist approach of phenotype uncertainty is very fast, and it is not computationally demanding.

The second sampling algorithm is the holdout sampler [6], which is inspired by the bootstrapping procedure, finds the minimum-scale genetic signature in each random holdout and the likelihood is established using the validation dataset using the k-NN mentioned above. Here, the prior probability distribution is established by the discriminatory power of the genetic signature in each random hold out and it is changing in each hold out. Therefore, this algorithm is based in a complete different sampling

paradigm than the FR sampler. This algorithm is also very fast and has as main advantage its robustness.

The third methodology concerns a random sampling (RS) of the set of genes that are differentially expressed in the phenotype. RS randomly selects genes within this set with a prior uniform distribution, building genetic signatures of different lengths (number of genes) within some bounds that are related to the phenotype problem complexity. Signatures that best predict the phenotype are considered for the posterior frequency analysis. This algorithm is also novel and it is introduced in this paper for comparative purposes.

Finally, the last procedure concerns the Bayesian Networks (BNs), a popular modeling formalism used in bioinformatics with many applications in modern genomics [7–10].

We show the application of these sampling methodologies to the analysis of the Triple Negative Breast Cancers (TNBC) phenotype, more precisely, the presented algorithms are used to predict samples with metastasis, unravelling the altered genetic pathways by the metastasis events. For that purpose, we used the microarray analyzed by Jézéquel et al. [11], which is deposited in the Gene Expression Omnibus (GEO) under the acronym GSE58812. This dataset contains the microarray analysis of 107 patients diagnosed with TNBC and controlled for metastasis (44 relapsed and 63 were disease-free after a follow-up period of 7 years), that were studied and treated between 1998 and 2007 at the Institut de Cancérologie de l'Ouest – René Gauducheau and the Institut de Cancérologie de l'Ouest – Paul Papin. This data received the consent of patients as required by the French Committee for the Protection of Human Subjects (CCPPRB). Gene expression analysis were performed in quality control RNA samples using Affymetrix® Human Genome U133 Plus 2.0 Arrays (Affymetrix®, Santa Clara, CA, USA), and the microarray analysis measured over 43,000 transcripts. Nevertheless, this comparison could also be applied to other diseases, using different types of genetic data. We have chosen this kind of cancers since it shows a high metastatic potential and very poor prognosis.

2 Sampling Algorithms

Genes Selection

Due to the large number of genetic probes that are monitored, a previous genes selection is realized for all the presented algorithms. For that purpose, a fold-change analysis is performed to find those genes that are differentially expressed in both tails (over and under-expressed) and reducing the sampling to this set. This way of acting tries to enhance the sampling of the header genes, which are those that expand the most important features of the phenotype prediction [5]. The rest are helper genes that explain high frequency details of the discrimination. The hypothesis is that mainly header genes are related to the genesis and the development of the disease.

It is important to understand that phenotype prediction problems have a huge uncertainty space, because the number of samples in this kind of experiments are much lower than the number of genetic probes that are monitored. This imbalance is even

greater due to the irruption of next generation DNA sequencing (NGS) technologies that serve to identify DNA bases while incorporating them into a nucleic acid chain. Each base emits a unique fluorescent signal as it is added to the growing strand, which is used to determine the order of the DNA sequence. NGS technology can be used to sequence the DNA from any organism, providing valuable information in response to almost any biological question. Also, DNA sequencing can be applied to small, targeted regions or the entire genome enabling researchers to investigate and better understand health and disease. Methods such as Whole-genome Sequencing, Target Resequencing and or Chip sequencing are currently used in this kind of research.

Fisher's Ratio Sampler (FRS)

The Fisher's ratio sampler takes into account the discriminatory power of the differentially expressed genes as measured by the Fisher's ratio to induce the prior sampling distribution of the high discriminatory genetic networks in a phenotype prediction problem. Figure 1 shows the Fisher ratio algorithm flowchart.

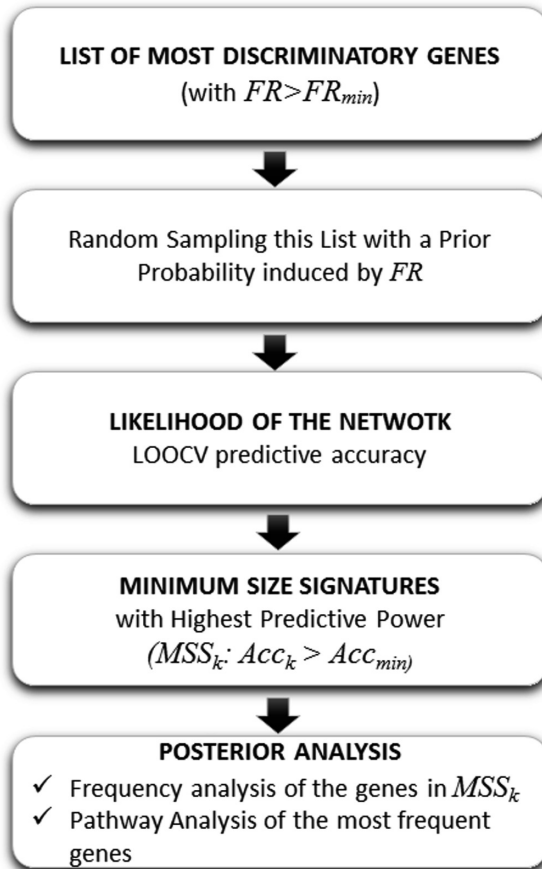


Fig. 1. Fisher's ratio sampler work-flow.

At the first step, this algorithm selects the most discriminatory genes having a Fisher's ratio greater than an established cut-off value [4]. Then, the algorithm finds the small-scale signatures that optimally discriminates between classes via recursive feature elimination. The predictive accuracy estimation is based on Leave-One-Out-Cross-Validation (LOOCV) and using a k-NN classifier [1]. The small-scale signature provides the typical length (number of genes) of the high discriminatory networks that are going to be sampled, and depends on the phenotype problem complexity. The greater the number of genes in this signature, the greater the complexity.

The next step is the random sampling of high discriminatory equivalent networks of genes of variable length, using a prior sampling probability of any individual gene proportional to its Fisher's ratio. The idea behind this method is to prioritize the discriminatory power of genes according to their Fisher's ratio. Genes are rated as *Headers* (variables with $FR > Tol$) and *Helpers* (variables with $FR < Tol$), and the Fisher sampler builds genes signatures at each step. Headers serve to expand the low frequency features in the discrimination, while helpers serve to span the high frequency details. Once a network is randomly built according to the Fisher's probability distribution, its LOOCV predictive accuracy is established. Finally, the posterior analysis of the most-frequently sampled genes in the most discriminatory networks, serves to establish the defective biological pathways.

The Holdout Sampler (HS)

The gene idea behind the hold out sampler is completely different from the Fisher's ratio sampler, but the aim is the same: sampling the uncertainty space which is intrinsic to phenotype prediction problems. For that purpose the idea consists in varying the evidence term $-P(\mathbf{c}^{obs})$ - in Bayes' expression:

$$P(\mathbf{g}/\mathbf{c}^{obs}) = P(\mathbf{g})L(\mathbf{c}^{obs}/\mathbf{g})/P(\mathbf{c}^{obs}) \quad (2)$$

The simplest way of doing that is performing random data bags with different datasets for training and blind validation. This is equivalent to change the evidence of \mathbf{c}^{obs} with respect to the given classifier \mathbf{L}^* , because some of the samples (in the validation set) have not been observed. This procedure, which is based in the idea of bootstrapping (or random sampling with replacement) proposed by [12]. This statistical technique is used to establish the confidence intervals in sample estimates and to estimate the sampling distribution of any statistic using random sampling methods. In this case, the idea has been adapted to sample the uncertainty space in phenotype prediction problems, based on the results shown in simple low-dimensional regression analysis [7], where it has been shown that a it is possible to efficiently sampling the posterior distribution of the model parameters by performing least-squares of different data bags.

The workflow of the holdout sampler is shown in Fig. 2. This algorithm samples the uncertainty space in two steps:

- (1) At the first step this algorithm generates different random 75/25 data bags (or holdouts), where 75% of the data in each bag is used for learning and 25% for blind validation. For each of these bags the genes are ranked according to their

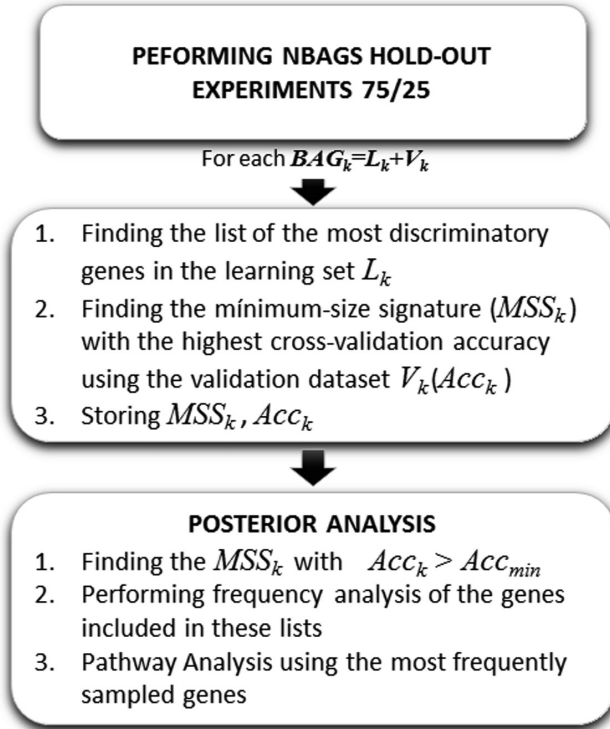


Fig. 2. Holdout sampler work-flow.

discriminatory power given by the Fisher's ratio, and the algorithm looks for the shortest list of prognostic variables with the highest predictive accuracy (minimum-scale genetic signature - MSS) via recursive feature elimination. The predictive accuracy estimation is based on Leave-One-Out-Cross-Validation (LOOCV) by averaging the LOOCV predictive accuracy over all the samples of the validation dataset in each bag. This methodology involves a k-Nearest-Neighbor classifier in the reduced set of high discriminatory genes (MSS) [1, 5].

- (2) The second step is the posterior analysis, which consists of finding all the minimum size signatures that have a validation predictive accuracy higher than a given tolerance. In this case, we have considered all the holdouts with predictive accuracy higher than 90%. This threshold is fixed high enough in order to provide explicative genetic networks of the phenotype. Nevertheless, it could be decreased till 80–85% if the number of high predictive networks is too small. Finally, we perform the frequency analysis of these lists to find the most-frequently sampled genes that serve to establish the defective genetic pathways [6].

Random Sampler (RS)

This novel algorithm randomly selects genes and builds signatures of variable length. This algorithm is similar to the Fisher's ratio sampler, but the prior sampling distribution is uniform within the set of differentially expressed genes, instead of being proportional to their Fisher's ratio. The predictive accuracy is established via LOOCV as in the FRS sampler. Signatures that best predict the phenotype are considered for the posterior frequency analysis. Finally, the most-frequently sampled genes are used to establish the genetic pathways and the biological processes involved.

The workflow of the Random Sampler (shown in Fig. 3) is similar to that of the Fisher's ratio sampler but the sampling prior probability is uniform in the set of discriminatory genes. This algorithm is obviously more exploratory than the Fisher's ratio sampler because it works with a less informative prior, but the main point is to show that the sampled networks should be similar mechanistically. This fact is a confirmation that we are sampling pathways and not individual genes that can have a given discriminatory power of the phenotype.

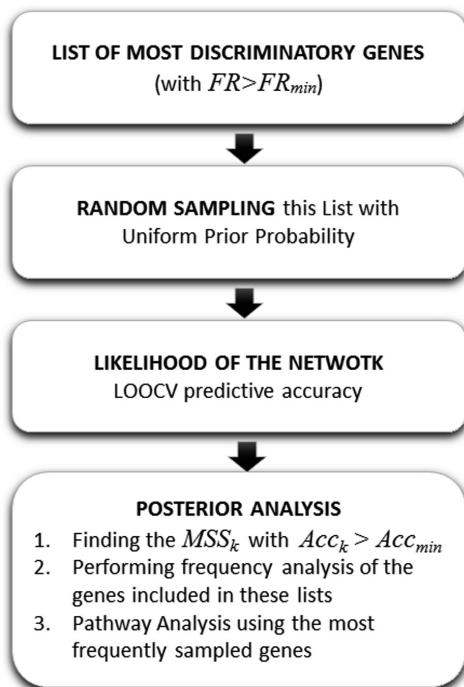


Fig. 3. Random sampler work-flow.

Bayesian Networks (BNs)

A Bayesian Network (BN) is a data structure that encodes the conditional probability distribution between variables by using a directed acyclic graph. This algorithm is used

to sample the posterior distribution of the genetic signatures, $P(\mathbf{g}/\mathbf{c}^{\text{obs}})$, according to Bayes' rule: $P(\mathbf{g}/\mathbf{c}^{\text{obs}}) \sim P(\mathbf{g})L(\mathbf{c}^{\text{obs}}/\mathbf{g})$, where $P(\mathbf{g})$ is the prior distribution to sample the genetic signatures and $L(\mathbf{c}^{\text{obs}}/\mathbf{g})$ is the likelihood of the genetic signature \mathbf{g} , that depends on the its predictive accuracy $O(\mathbf{g})$.

In practice, this algorithm performs genes selection at two steps. First, the network structure is learned from the training dataset. To determine the best network model, the algorithm computes the marginal likelihood of candidate network structures, conditioned upon the data, and then chooses the network model that maximizes the marginal likelihood. The learned network illustrates how functionally genes influence each other and also serves as a predictor of the phenotype. At the second step, the parameters of this network are trained by looking at the conditional probabilities of the variables within the dataset. Finally, the phenotype prediction is performed using a variable elimination algorithm, as described in [13]. The genes associated to the final BN are used for the detective pathways identification [8, 9, 11]. To accomplish the Bayesian Networks modelling we have used CGBayesNets [10]. CGBayesNets looks for the network that maximizes the marginal likelihood of the data. Since the number of possible networks grows exponentially with the number of gene candidates, all networks cannot be investigated, and different heuristics are employed in the search for good networks.

The workflow of this algorithm is shown in Fig. 4.

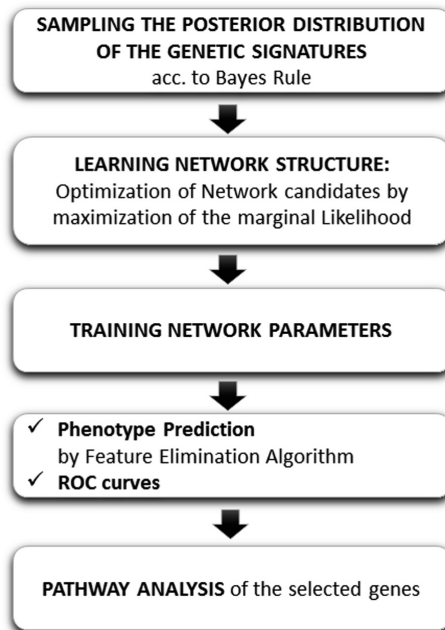


Fig. 4. Bayesian Network sampler work-flow.

The fact that CGBayesNets uses different search techniques to optimize the discriminatory networks makes that this algorithm is not a pure sampling algorithm. In fact, for pathway analysis it would be much better to sample the uncertainty space of the high discriminatory Bayesian Networks, because this set is not a unique representation of the phenotype problem uncertainty, that is, other plausible networks might exist that explain the phenotype with a similar likelihood.

Sampling the set of these networks constitutes a subject for further research. Besides, the noise in data and class assignment [2] falsifies the pathways analysis and impacts the BNs search and optimization. That way, inferring the genetic pathways via the genes involved in the best network might not be robust enough.

Defective Genetic Pathways Identification

The defective genetic pathways and biological processes involved were established using GeneAnalytics [14]. This software uses the main ontological databases (Biosystems, Reactome, Qiagen, Kegg, Cell Signaling Technology and R&D Systems), and provides important clues about the biological processes involved, and the existence of chemical compounds to target the actionable genes characterizing the phenotype. In all the cases, this analysis is performed over the high discriminatory networks of genes, provided by each prediction algorithm.

3 Results

Table 1 shows the list of the 15 first genes most frequently sampled by each algorithm. It can be observed that the most frequent gene LINC00630 is the same in both Fisher's ratio sampler and Random sampler algorithms, and it is also sampled by Holdout

Table 1. List of most-frequently sampled genes by the different algorithms.

Fisher's ratio sampler	Hold out sampler	Random sampler	Bayesian network
LINC00630	OTUB2	LINC00630	ZNF597
LOC100506272	STC1	HIPK3	ZDHHC2
STC1	BAIAP2-AS1	CCDC116	YY1
BAIAP2-AS1	STC1	EXOC5	SPP1
ARFGAP2	KCNS2	GHSR	SMAD9
LHX9	LOC100506272	ZNF540	SHANK1
LOC646482	LOC644135	ATF3	RBMS3
CACNA1S	LINC00630	1557882_at	PRICKLE1
AC108056.1	UGT1A1	220899_at	PRDM11
NXF3	ARFGAP2	ARFGAP2	PML
GIPC3	CACNA1I	CXADR	NAV1
KCNS2	DCAF8	AHI1	MASP1
DAZ1	RP11-799D4.4	KIRREL3-AS3	LOC646482
UGT1A1	MDM2	DRP2	LOC101927735
RP5-855D21.1	RP11-38C18.3	207743_at	P4HA2
EXOC5	BFSP2-AS1	JMJD6	LINC00642

is involved in gene expression and also in transcription. ZDHHC2 is called Palmitoyltransferase or Reduced expression associated with metastasis protein. It has been associated to human colorectal cancers with liver metastasis [16]. Yin Yang 1 (YY1) is highly expressed in various types of cancers and regulates tumorigenesis through multiple pathways. YY1 is generally overexpressed in breast cancer cells [17]. This provides an idea of the importance of the genes involved in the BNs.

Table 2 shows the list of the most frequently pathways sampled by each algorithm with their relative score. The main conclusions obtained from this analysis are the following:

Table 2. Top score pathways sampled by the different algorithms.

Fisher's ratio sampler		Holdout sampler	
Score	Top pathways	Score	Top pathways
10.3	Direct P53 effectors	11.2	JNK signaling in the CD4 + TCR pathway
10.1	DREAM repression and dynorphin expression	9.7	RhoA signaling pathway
9.6	P53 signaling	8.3	ATM pathway
8.8	RhoA signaling pathway	8.1	FoxO signaling pathway
8.4	P53 pathway (RnD)	8.0	TGF-beta signaling pathway
Random sampler		Bayesian network	
Score	Top pathways	Score	Top pathways
15.0	DREAM repression and dynorphin expression	8.1	Direct P53 effectors
10.1	Direct P53 effectors	8.0	Proteolysis putative SUMO-1 pathway
10.1	Immune response role of DAP12 receptors in NK Cells	7.1	Creation of C4 and C2 activators
9.9	JNK signaling in the CD4 + TCR pathway	6.9	TGF-beta receptor signaling
9.8	MAPK signaling pathway	6.8	MTOR signaling pathway

1. Direct P53 Effectors was identified by three of the four presented algorithms. Fisher ratio, Random Sampler and Bayesian Network commonly identified Direct P53 Effectors pathway. Other common pathways are DREAM Repression and Dynorphin Expression (between Fisher's ratio and Random sampler) and TGF-beta Signaling Pathway (between Hold-out sampler and Bayesian Network). Additionally, Bayesian Network also share with Fisher's ratio P53 Signaling and P53 Pathway (RnD), with a lower score.
2. The Bayesian network produces the list with fewer discriminatory genes (only 68). This fact impacts the pathway identification. Nevertheless, some of the pathways are related to the other samplers: The Direct P53 effectors and the TGF-beta receptor signaling pathways. The main biological process involved is the Complement Activation via the Lectin pathway.

3. The random sampler is very fast and can be run many iterations to unravel different sets of discriminatory genes. This algorithm found 410 different genetic probes with a sampling frequency higher than 0.15%. The main pathways found are: Dream repression, Direct P53 effectors, Immune response, JNK signaling and MAPK signaling. These pathways are very similar to those found by the Holdout sampler [6].
4. The pathways and biological identified by the Bayesian networks have a lower score than those found by the random sampler, Fisher's ratio [4] and Holdout samplers [6].
5. In all the cases, the pathways are involved in cancer and also in immune response. Therefore, this analysis confirms the insight that the altered pathways should be independent of the sampling methodology and the classifier that is used to infer them. This result should be confirmed analyzing other diseases and using different kind of classifiers.

4 Conclusions

In this paper, we have compared different gene samplers in phenotype prediction problems, the random sampler, the Fisher's ratio sampler and the Holdout sampler, to the Bayesian Networks. Among all of them, the Bayesian networks is the less efficient due to the much high computational effort needed to find the discriminatory networks.

The random sampler provides a good result because it very fast and it is performed in the set of differentially expressed genes. This analysis shows common pathways related to cancer progression and immune response. This preliminary analysis shows that the Fisher's ratio, the holdout and the random samplers are a good alternative to Bayesian networks, since they are much faster, and they need a lower computational effort because they do not need to infer the posterior probability distribution. This result could be improved by considering other possible (or equivalent) Bayesian Networks with lower predictive accuracy, since this network (factorization) is not unique. Further research will be devoted in the future to analyze this important subject.

References

1. De Andrés Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Design of biomedical robots for phenotype prediction problems. *J. Comput. Biol.* **23**(8), 678–692 (2016)
2. De Andrés-Galiana, E.J., Fernández-Martínez, J.L., Sonis, S.: Sensitivity analysis of gene ranking methods in phenotype prediction. *J. Biomed. Inf.* **64**, 255–264 (2016)
3. Fernández-Martínez, J.L., Fernández-Muñiz, M.Z., Tompkins, M.J.: On the topography of the cost functional in linear and nonlinear inverse problems. *Geophysics* **77**(1), W1–W15 (2012). <https://doi.org/10.1190/geo2011-0341.1>
4. Cernea, A., Fernández-Martínez, J.L., deAndrés-Galiana, E.J., Fernández-Ovies, F.J., Fernández-Muñiz, Z., Álvarez-Machancoses, O., Saligan, L.N., Sonis, S.: Sampling defective pathways in phenotype prediction problems via the Fisher's ratio sampler. In: *IWBIO 2018* (2018)

5. Saligan, L.N., Fernández-Martínez, J.L., de Andrés Galiana, E.J., Sonis, S.: Supervised classification by filter methods and recursive feature elimination predicts risk of radiotherapy-related fatigue in patients with prostate cancer. *Cancer Inf.* **13**(141–152), 2014 (2014)
6. Fernández-Martínez, J.L., Cernea, A., de Andrés-Galiana, E.J., Fernández-Ovies, F.J., Fernández-Muñiz, Z., Álvarez-Machancoses, O., Saligan, L.N., Sonis, S.: Sampling defective pathways in phenotype prediction problems via the Holdout sampler. In: *IWBBIO 2018* (2018)
7. Jiang, X., Barmada, M.M., Visweswaran, S.: Identifying genetic interactions in genome-wide data using Bayesian networks. *Genet. Epidemiol.* **34**(6), 575–581 (2010)
8. Hageman, R.S., Leduc, M.S., Korstanje, R., Paigen, B., Churchill, G.A.: A Bayesian framework for inference of the genotype-phenotype map for segregating populations. *Genetics* **187**(4), 1163–1170 (2011)
9. McGeachie, M.J., Chang, H.H., Weiss, S.T.: CGBayesNets: conditional gaussian Bayesian network learning and inference with mixed discrete and continuous data. *PLoS Comput. Biol.* **10**(6), e1003676 (2014)
10. Su, C., Andrew, A., Karagas, M.R., Borsuk, M.E.: Using Bayesian networks to discover relations between genes, environment, and disease. *BioData Mining* **6**, 6 (2013)
11. Jézéquel, P., Loussouarn, D., Guérin-Charbonnel, C., Campion, L., et al.: Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* **20**(17), 43 (2015)
12. Efron, B., Tibshirani, R.: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, Boca Raton (1993). ISBN 0-412-04231-2
13. Koller, D., Friedman, N.: *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, Cambridge. xxxv, 1231 p. (2009)
14. Stelzer, G., Inger, A., Olender, T., Iny-Stein, T., Dalah, I., Harel, A., et al.: GeneDecks: paralog hunting and gene-set distillation with GeneCards annotation. *OMICS* **13**(6), 477 (2009)
15. Qin, N., Wang, C., Lu, Q., et al.: A cis-eQTL genetic variant of the cancer–testis gene *CCDC116* is associated with risk of multiple cancers. *Hum. Genet.* **136**, 987 (2017). <https://doi.org/10.1007/s00439-017-1827-2>
16. Oyama, T., Miyoshi, Y., Koyama, K., Nakagawa, H., Yamori, T., Ito, T., Matsuda, H., Arakawa, H., Nakamura, Y.: Isolation of a novel gene on 8p21. 3–22 whose expression is reduced significantly in human colorectal cancers with liver metastasis. *Genes Chromosomes. Cancer* **29**, 9–15 (2000)
17. Wan, M., Huang, W., Kute, T.E., Miller, L.D., Zhang, Q., Hatcher, H., Wang, J., Stovall, D. B., Russell, G.B., Cao, P.D., Deng, Z., Wang, W., Zhang, Q., Lei, M., Torti, S.V., Akman, S.A., Sui, G.: Yin Yang 1 plays an essential role in breast cancer and negatively regulates p27. *Am. J. Pathol.* **180**(5), 2120–2133 (2012). <https://doi.org/10.1016/j.ajpath.2012.01.037>

Biomedical Engineering



Composite Piezoelectric Material for Biomedical Micro Hydraulic System

Arvydas Palevicius^(✉), Giedrius Janusas, Elingas Cekas,
and YatinkumarRajeshbhai Patel

Faculty of Mechanical Engineering and Design,
Kaunas University of Technology, Studentu str. 56, 51424 Kaunas, Lithuania
arvydas.palevicius@ktu.lt

Abstract. ‘Lab-on-a-chip’ is integrated micro-analytical system, which could perform sample pre-treatment, chemical reactions, analytical separation, detection and data handling. These platforms are able to convert biological, chemical or mechanical responses into electrical signals using the piezoelectric or piezoresistive materials. This paper discusses a piezoelectric composite material displaying its mechanical properties such as resonant frequencies, Young’s modulus and density. Nano composite polymer highlights the property of piezo effect and is suitable for formation of periodic micro scale patterns on it. These micro patterns are intended to be used as innovative functional elements in biomedical micro hydro mechanical systems such as micro channels. Thus by controlling surface configuration and the shape of active deformable polymer, pressure in microfluidic vessels can be changed and mobility of the transported bioparticles can be ensured.

Keywords: Piezoelectric composite · Micro hydrolic system · Microchanel
Bioparticals

1 Introduction

Miniaturization of analytical instruments utilizing microfabrication technology has increased the broad usage of analytical chemistry. Main focus was an increasing demand for low-cost instruments able to analyze compounds of very small volume with a high level of automation and precision. The terms ‘Micro-Total Analysis Systems (μ -TAS)’, and also ‘Lab-on-a-chip’, aims to develop integrated micro-analytical systems [1], which could complete analysis cycles (e.g. sample pre-treatment, chemical reactions, analytical separation, detection, and data handling) on the same micro device [2]. In addition to multi-functional abilities, micro fabricated analytical systems should enhance performance, like fast response and increased analysis speed [3].

One of the recent new applications in this area is related to bio sensing elements based on cantilever or membrane type sensing platforms. These platforms are able to convert biological responses into electrical signals [4] using the piezo effect created by the piezoelectric material, on which the active element of the system is based.

Many various materials are used as an active element in the sensing platforms, whether it is a micro-pump, a biosensor or a drug delivery device. PDMS

(Polydimethylsiloxane) [5, 6], PU (Polyurethane) [5], PZT (Lead zirconate titanate) [7], Si (Silicon) [8], and SiO (Silicon oxide) [9] have shown a great promise, but all of them lack of flexibility and abilities to change their parameters on-demand. To increase the sensitivity and ability to change parameters of the material a high Q factor is needed [10]. This paper discusses a piezoelectric composite material displaying its mechanical properties such as resonant frequencies, Young's modulus and density. Nano composite polymer highlights the property of piezo effect and is suitable for formation of periodic micro scale patterns on it. These micro patterns are intended to be used as innovative functional elements in biomedical micro hydro mechanical systems such as micro channels. Thus by controlling surface configuration and the shape of active deformable polymer, pressure in microfluidic vessels can be changed and mobility of the transported bioparticles can be ensured.

2 Experimental Setup

2.1 Piezocomposite Synthesis

Piezocomposite material used in experiments was created using PZT nanopowder and three different binding polymers: polyvinyl butyral (PVB), polymethyl methacrylate (PMMA) and polystyrene (PS). 80% of PZT nanopowder was mixed with solution of polymer in benzyl alcohol under defined conditions: 80% of PZT and 20% of binding material [11]. Thus, three coatings of different thickness were screen printed and investigated: element 1 with a PZT + PVB coating of 50 μm thickness, element 2 with PZT + PMMA coating of 60 μm thickness, and element 3 with PZT + PS coating of 60 μm thickness. More details about the synthesis of piezocomposite are presented in paper [11]. Afterwards, a laminar microstructure of 4 μm period could be embedded into formed piezoelectric film using the thermal embossing technology. Moreover, it is possible to increase optical parameters (surface Plasmon resonance) and ensure antibacterial properties of the formed microstructures using silver nanoparticles. Silver nanoparticles were formed from the solution of 0.05 AgNO_3 in deionized water and dip-coated on top of the periodic microstructure [12].

2.2 Mechanical Material Properties

The properties of the binding materials are presented in Table 1. Using these properties, the density of every used composite material was calculated. The density of PZT + PVB, PZT + PMMA, and PZT + PS were 6294 kg/m^3 , 6314 kg/m^3 , and 6290 kg/m^3 respectively.

Table 1. PVB, PMMA, and PS properties.

Property	PVB	PMMA	PS
Tensile strength	≥ 20 MPa	48–76 MPa	32–44 MPa
Poisson's ratio	0.45–0.49	0.35–0.4	0.4–0.41
Density	1070 kg/m^3	1170 kg/m^3	1050 kg/m^3

The bottom layer of multilayer structure is a copper foil made from UNS C10100 copper alloy. Mechanical properties of used copper alloy are presented in Table 2.

Table 2. Mechanical properties of UNS C10100 copper alloy.

Property	Value
Hardness, Vickers (1/2)	75.0–90.0
Tensile strength	221–455 MPa
Elongation at break	55%
Modulus of elasticity	115 GPa
Poisson's ratio	0.31
Machinability	20%
Shear modulus	44.0 GPa

2.3 Multilayer Structure

A cantilever type multilayer elements were created. The proposed element consisted of two layers: copper foil and composite piezoelectric material. Geometrical representation of multilayer structures is presented in Fig. 1.

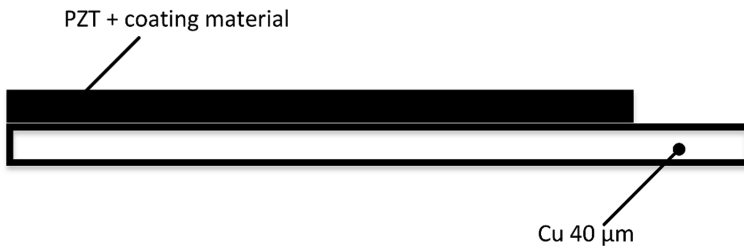


Fig. 1. Schematical view of the multilayer specimen.

Geometrical dimensions of these three elements are given in Table 3.

Table 3. Specimen dimensions

Specimen	Coating length (C), mm	Clamped length (L), mm	Width (W), mm	Total thickness, nm	Coating thickness, mm
PVB	15	32	18	90	50
PMMS	17	30	15	100	60
PS	20	33	21	100	60

2.4 Dynamic Investigation of Composite Piezoelectric Material

Damping ratio and modulus of elasticity of the piezocomposite materials were determined using hybrid experimental-numerical method. Dynamical response of the multilayer cantilever in to impulse excitation were measured using laser triangular displacement sensor LK-G3000 series together with sensor head of LK-G82 and control block LK-G3001PV made by Keyence from Illinois, US. Signal from control block and the generated voltage were collected with data acquisition system, a 4 channel USB oscilloscope PICO 3424, and forwarded to a computer. Obtained data was then analyzed with Picoscope 6 software. Experimental setup scheme used to measure dynamic characteristics of piezoelectric elements is presented in Fig. 2.

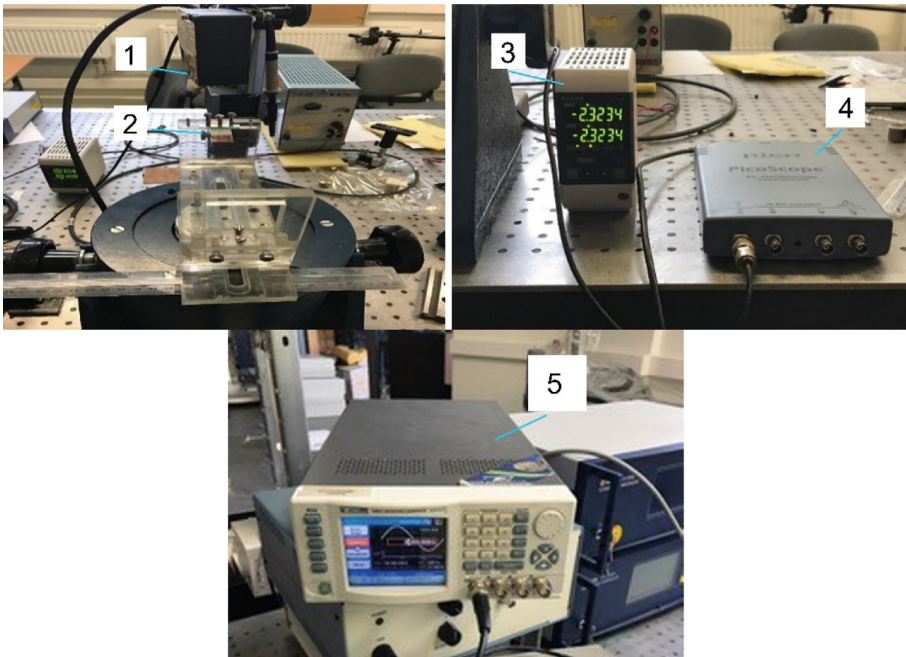


Fig. 2. Experimental setup for response characteristic measurements of multilayer piezoelectric element consist of: (1) sensor head LK-G82, (2) horizontally clamped specimen, (3) control block LK-G3001 PV, (4) Pico Scope oscilloscope, (5) power and wave form generator.

2.5 Theoretical Calculations

In order to compare the experimental and theoretical models, the natural frequency of proposed multilayer specimen had to be calculated by the following formulas 1–3. Logarithmic decrement δ from experimental data for n observed periods was calculated according to formula:

$$\delta = \frac{1}{n} \ln \frac{x(t)}{x(t+nT)}, \quad (1)$$

where, $x(t)$ – amplitude at time t and $x(t+nT)$ – amplitude at time $t+nT$ (after n periods). Then damping ration ξ was calculated from formula:

$$\xi = \frac{1}{1 + \left(\frac{2\pi}{\delta}\right)^2}. \quad (2)$$

Finally the natural frequency ω_n of the multilayer structure was calculated according to formula:

$$\omega_d = \omega_n \sqrt{1 - \xi^2}, \quad (3)$$

where ω_d – the damped natural frequency calculated from experimental data.

3 Results and Discussion

Measured response of multilayer structure with PVB binding materials into pulse excitation is presented in Fig. 3. Measured natural frequency of 31.5 Hz were calculated using formulas 1–3. From numerical simulations natural frequency of 32.1 Hz was determined (Fig. 4) for multilayer construction with PVB binding material, when Young's modulus is 3.9 GPa. It is only 1.5% difference in natural frequencies. It could be stated that Young's modulus of PZT + PVB composite material is 3.9 GPa.

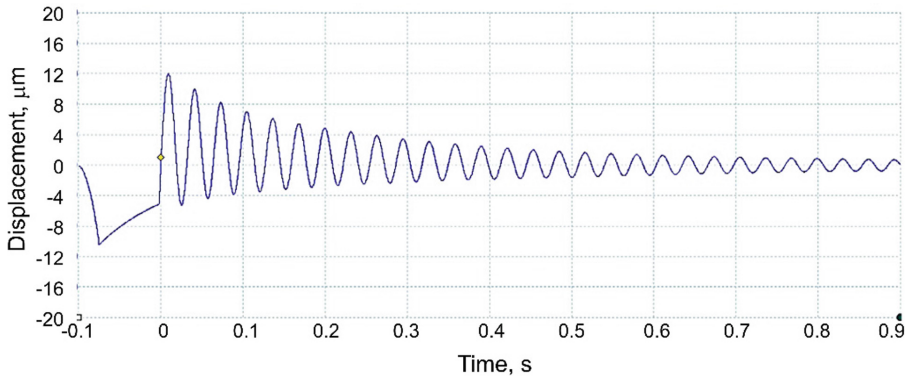


Fig. 3. Graphical representation of the logarithmic decrement of the specimen with PVB binding material.

The experimentally measured natural frequency of specimen with PS binding material was 43.15 Hz (Fig. 5). The simulation did verify this value as the model gave the result of resonant frequency of 43.02 Hz (Fig. 6) with the accuracy of 0.3% when Young's modulus of PZT + PS composite material is equal to 5.3 GPa.

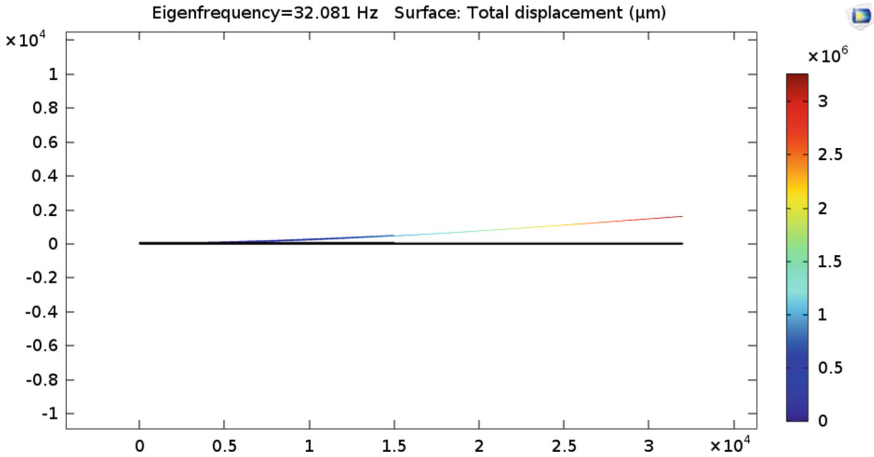


Fig. 4. Resonant frequency of the specimen with PMMA binding material.

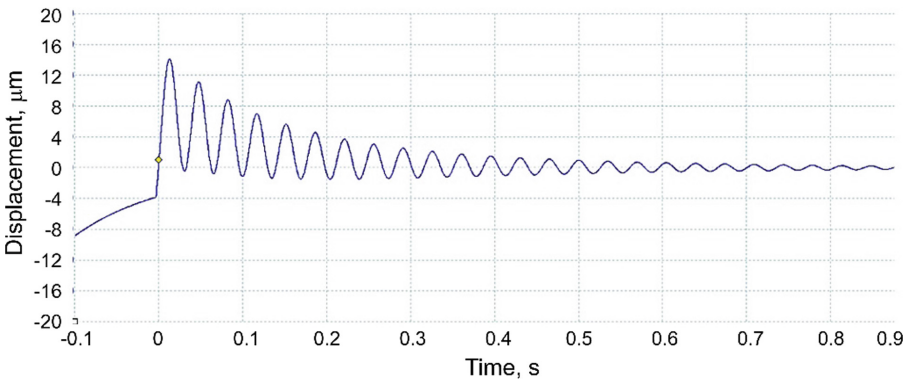


Fig. 5. Graphical representation of the logarithmic decrement of the specimen with PMMA binding material.

The PZT + PMMA composite material has shown the highest Young’s modulus of all three specimens. The value of the Young’s modulus for PMMA is 6.3 GPa. It was also calculated comparing the theoretical (Fig. 8) and experimental (Fig. 7) results of the specimen with PMMA binding material. Theoretical natural frequency was 46.99 Hz and experimental value – 46.17 Hz. Only 1.8% difference has been noted.

The presented material could be used for formation of micro channels with controllable parameters using thermal replication process. Numerically behavior of the system of micro channels of $L = 20 \mu\text{m}$ length, $P = 4 \mu\text{m}$ period, $T = 2 \mu\text{m}$ thickness, $d = 1 \mu\text{m}$ depth, $w = 2 \mu\text{m}$ width of the land and $2r = 2 \mu\text{m}$ width of the ridge (Fig. 9a) was investigated using COMSOL Multiphysics 5.2a software. It is a periodical system, therefore just one microchannel was analyzed in the model with the symmetrical boundary conditions (Fig. 9b) i.e. left and right surfaces were fixed using rollers.

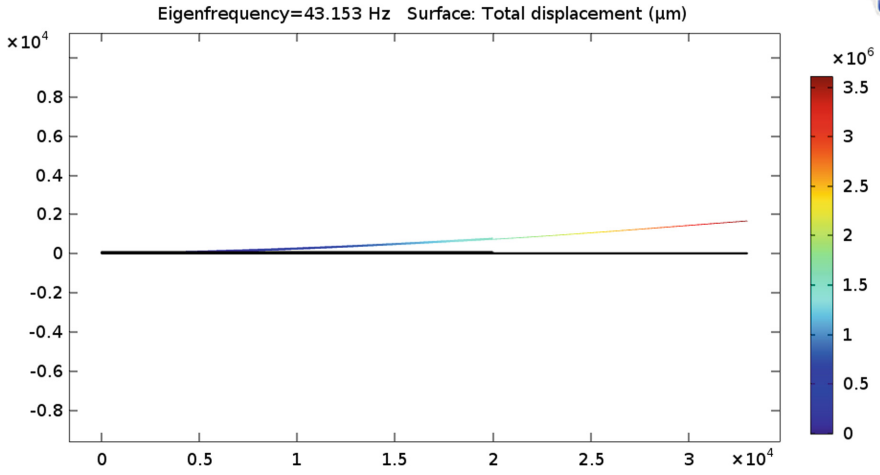


Fig. 6. Resonant frequency of the specimen with PMMA binding material.

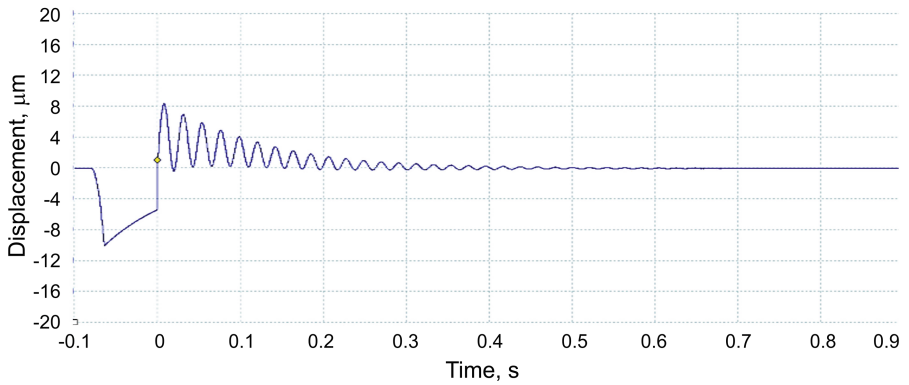


Fig. 7. Graphical representation of the logarithmic decrement of the specimen with PS binding material.

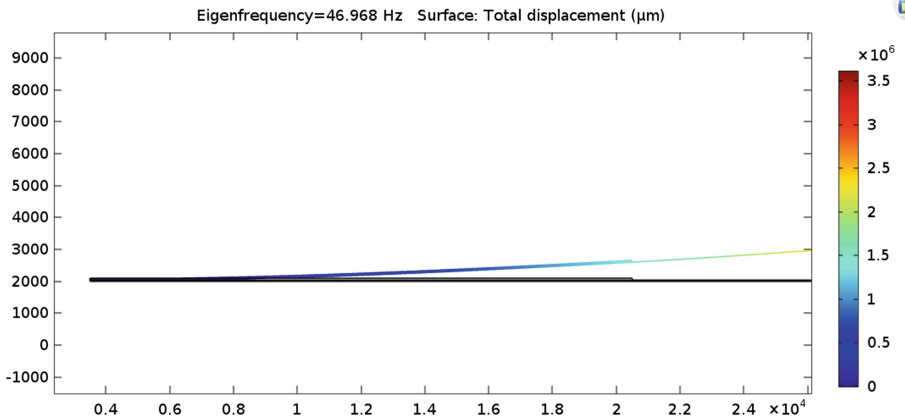


Fig. 8. Resonant frequency of the specimen with PS binding material.

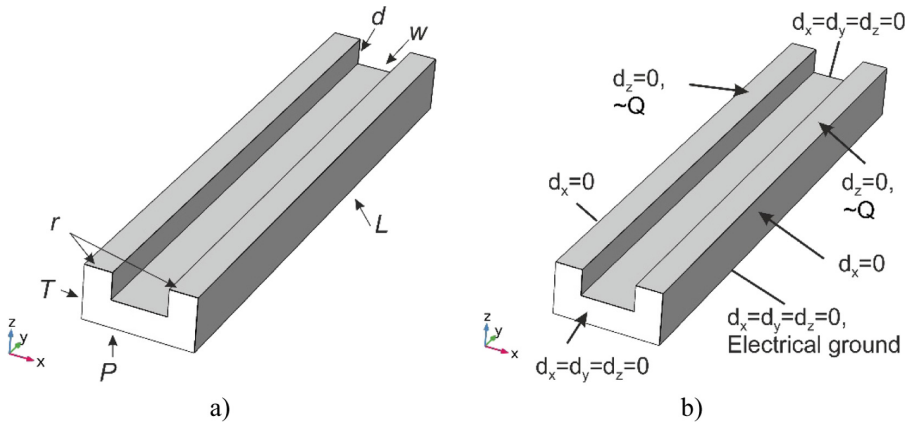


Fig. 9. Dimensions (a) and boundary conditions (b) of the finite element model of microchannel.

The front and back surfaces were fixed totally because it is attached to fluid containers. The bottom surface will be attached to other components of the sensor, therefore it is totally fixed and it serves as an electrical ground. The top of microchannel system is covered by electrode and rigid transparent layer which allows for the visual inspection of the fluid flow and ensures the high pressure inside of the channel. The system of micro channels is excited by sinusoidal alternating electrical potential of 20 V.

Three examples of the dynamical response of the suggested microchannel are presented in Figs. 10 and 11. The total deformation and X component of the displacement field of the electrically excited system at 110 MHz frequency shows that it works as a mechanical valve i.e. the cross-section area increases or reduces evenly

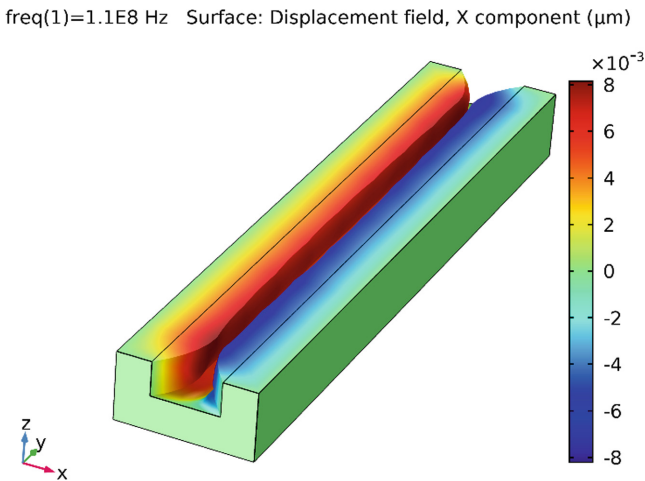


Fig. 10. Total deformation of periodically at 110 MHz frequency excited microchannel with visualized X component of the displacement field.

through all the length of microchannel (Fig. 10). While increased frequency to 122 MHz allows to divide channel into two segments with different concentrations of bioparticles (Fig. 11).

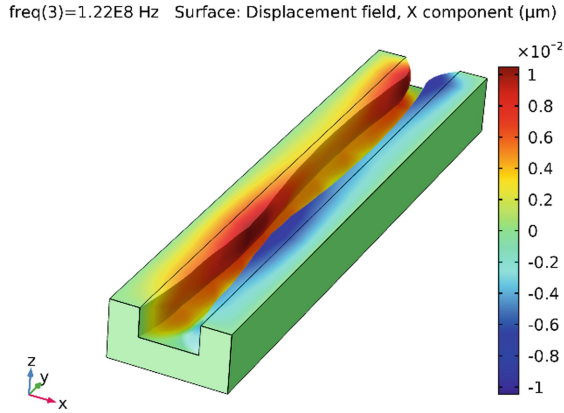


Fig. 11. Total deformation of periodically at 122 MHz frequency excited microchannel with visualized X component of the displacement field.

An example of bioparticles transportation in microchannel is presented in Fig. 12.

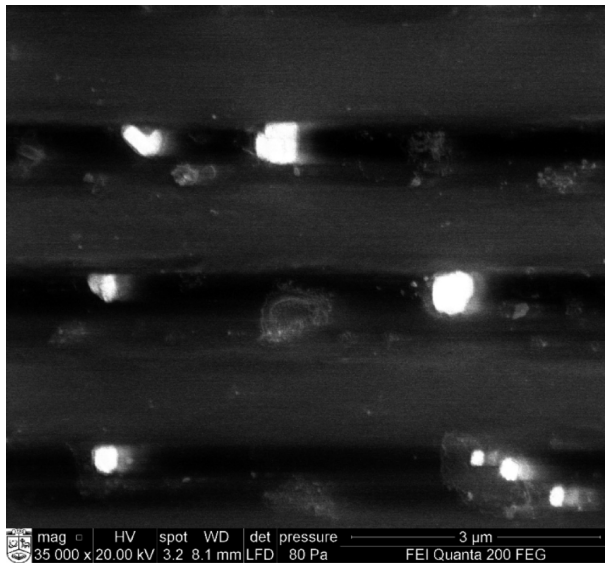


Fig. 12. SEM view of bioparticles transportation.

4 Conclusions

New type composite piezoelectric material integrated to an actuator was proposed. Experimental results have shown the natural frequencies of PVB, PMMA, and PS composite piezoelectric material are 32.1 Hz, 46.99 Hz, and 43.15 Hz respectively. Due to a very low difference between theoretical and experimental values (max. deviation of 1.8%) Young's modulus can be clearly stated for these composite materials. Young's modulus of PZT mixed with PVB value is 3.9 GPa, PMMA – 6.3 GPa, and PS – 5.3 GPa. A high Q factor of the proposed composite materials would allow to easily control the micro device as micro channels. The ability to change active materials parameters during operation would increase the range of usage and sensing abilities of the system.

Acknowledgments. This research was funded by a grant S-MIP-17-102 from the Research Council of Lithuania.

References

1. Burns, M.A., Johnson, B.N., Brahmasandra, S.N., Handique, K., Webster, J.R., Krishnan, M., Sammarco, T.S., Man, F.P., Jones, D., Heldsinger, D., Mastrangelo, C.H., Burke, D.T.: An integrated nanoliter DNA analysis device. *Science* **282**, 484–490 (1998)
2. Manz, A., Graber, N., Widmer, H.M.: Miniaturized total chemical-analysis systems—a novel concept for chemical sensing. *Sens. Actuators B* **1**, 244–248 (1990)
3. Jakobson, S.C., Culbertson, C.T., Daler, J.E., Ramsey, J.M.: Microchip structures for submillisecond electrophoresis. *Anal. Chem.* **70**, 3476–3480 (1998)
4. Ziegler, C.: Cantilever-based biosensors. *Anal. Bioanal. Chem.* **379**, 946–959 (2004)
5. Arefin, A., Huang, J.H., Platts, D., et al.: Fabrication of flexible thin polyurethane membrane for tissue engineering applications. *Biomed. Microdevices* **19**, 98 (2017)
6. Holley, M.T., YekrangSafakar, A., Maziveyi, M., et al.: Measurement of cell traction force with a thin film PDMS cantilever. *Biomed. Microdevices* **19**, 97 (2017)
7. Ogawa, J., Kanno, O., Kotera, H., Wasa, K., Suzuki, T.: Development of liquid pumping devices using vibrating microchannel walls. *Sens. Actuators A: Phys.* **152**, 2 (2009)
8. Medjahdi, N., Benmoussa, N., Benyoucef, B.: Modeling, simulation and optimization of the mechanical response of micromechanical silicon cantilever: application to piezoresistive force sensor. *Phys. Proced.* **55**, 348–355 (2014)
9. Stolyarova, S., Cherian, S., Raiteri, R., Zeravik, J., Skladal, P., Nemirovsky, Y.: Composite porous silicon-crystalline silicon cantilevers for enhanced biosensing. *Sens. Actuators B Chem.* **131**, 509–515 (2008)
10. Lu, J., Ikehara, T., Zhang, Y., Mihara, T., Itoh, T., Maeda, R.: High quality factor silicon cantilever driven by PZT actuator for resonant based mass detection. In: *Dans Symposium on Design, Test, Integration and Packaging of MEM/MOEMS – DTIP 2008*, Nice, France (2008)
11. Janusas, G., Ponelyte, S., Brunius, A., Guobiene, A., Vilkauskas, A., Palevicius, A.: Influence of PZT coating thickness and electrical pole alignment on microresonator properties. *Sensors* **16**(11), 1–9 (2016)
12. Ponelyte, S., Palevicius, A.: Novel piezoelectric effect and surface plasmon resonance-based elements for MEMS applications. *Sensors* **14**, 6910–6921 (2014)



Trabecular Bone Score in Overweight and Normal-Weight Young Women

Abdel-Jalil Berro⁶, Marie-Louise Ayoub^{1,2,3}, Antonio Pinti⁵✉, Said Ahmaidi⁶, Georges El Khoury^{1,4}, César El Khoury^{1,5}, Eddy Zakhem¹, Bernard Cortet^{2,3}, and Rawad El Hage¹

¹ Department of Physical Education, Faculty of Arts and Social Sciences, University of Balamand, El-Koura, Lebanon

² University Lille (ULCO, USTL, Lille 2), Lille, France

³ Physiopathology of Inflammatory Bone Diseases, EA 4490, Boulogne sur Mer, France

⁴ Laboratoire VIP'S, UFR-APS, Campus la Harpe, Université Rennes 2, Rennes, France

⁵ I3MTO, EA 4708, Université d'Orléans,

1, Rue Porte-Madeleine, BP 2439, 45032 Orléans, Cedex 1, France

antonio.pinti@univ-valenciennes.fr

⁶ Université de Picardie Jules Verne,

Allée Paschal-Grousset, 80025 Amiens, Cedex 1, France

Abstract. The aim of this study was to compare Trabecular Bone Score (TBS) in overweight and normal-weight young women. This study included 14 overweight (BMI > 25 kg/m²) and 42 normal-weight (BMI < 25 kg/m²) young Lebanese women whose ages range from 18 to 32 years. Body composition, Bone Mineral Content (BMC), Bone Mineral Density (BMD), and lumbar spine (L1–L4) TBS were assessed by dual-energy X-ray absorptiometry (DXA). The DXA measurements were completed for the whole body (WB), the lumbar spine (L1–L4), the total hip (TH) and the femoral neck (FN). Physical activity, daily calcium intake, daily protein intake and sleep quality index were evaluated using validated questionnaires. Maximal oxygen consumption (VO₂ max in l/min) was measured whilst exercising on a bicycle ergometer using a specialized device. Weight, height, BMI, lean mass, fat mass, WB BMC, WB BMD, TH BMD and FN BMD were significantly higher in overweight women compared to normal-weight women. Trabecular Bone Score (TBS) was not significantly different between the two groups (overweight and normal-weight). In the whole population (n = 56), weight, height, BMI, lean mass and fat mass were positively correlated to BMC and BMD values but not to TBS values. VO₂ max (l/min) was positively correlated to BMC, BMD and TBS (p < 0.05). This study suggests that being overweight is not associated with higher trabecular bone score values in young women.

Keywords: DXA · Body composition · Bone microarchitecture
Peak bone mass

1 Introduction

Osteoporosis is known to mainly affect postmenopausal women [1–3]. Structural adaptations to bone loss in aging women have been previously described [2]. Bone strength is influenced by several fractures such as BMD, bone geometry, cortical porosity and bone microarchitecture [4–6]. Bone microarchitecture is one of the factors that directly influences fracture risk [4–6]. The trabecular bone score (TBS) is a texture parameter that evaluates pixel gray level variations in DXA images of the lumbar spine [4–17]. The TBS variations may reflect bone microarchitecture [4–17]. Recent studies have shown that TBS is related to fractures in elderly women [9–12, 15–17]. We have previously shown that obesity and overweight are associated with higher BMD values at the hip and the lumbar spine in women [3, 18–21]. However, we have also shown that the correlation between BMD and TBS is relatively low suggesting that BMD and TBS reflect different bone properties and may not have the same predictors [22, 23]. Moreover, little is known concerning the influence of morphological characteristics (weight, height, BMI, lean mass and fat mass) on TBS in young adult women.

2 Purpose

The main aim of this study was to compare trabecular bone score in overweight and normal-weight young Lebanese women. We have hypothesized that the influence of being overweight on TBS values would be different from that on BMD.

3 Materials and Methods

3.1 Subjects and Study Design

56 young women whose ages range from 18 to 32 years participated in this study. The study participants were recruited from 2 private universities located in North Lebanon. All participants were nonsmokers and had no history of major orthopedic problems or other disorders known to affect bone metabolism. Other inclusion criteria included no diagnosis of comorbidities and no history of fracture. Moreover, women participating in this study were not pregnant and had not taken hormonal contraceptives for the past 6 months. Participants were divided into two groups (14 overweight (BMI > 25 kg/m²) and 42 normal-weight (BMI < 25 kg/m²) women) using international cut-offs for body mass index (BMI) [24]. This study was approved by the University of Balamand Ethics Committee.

3.2 Anthropometrical Measurements

Weight and height were measured, and body mass index (BMI) was calculated. Body composition was assessed by DXA (GE-Lunar, iDXA). In our laboratory, the in vivo coefficients of variation were <1% for fat and lean mass [21–23, 25].

3.3 VO₂ Max Assessment

We directly assessed VO₂ max of the participants using a Cosmed Fitmate pro device (version 2.20) while exercising on a bicycle ergometer (Siemens-Elema RE 820; Rodby Elektronik AB, Enhorna, Sweden). A progressive 2-minstep protocol (20–30 W/step) was used as previously described [25].

3.4 Physical Activity

The duration of physical activity (h/wk) was evaluated using a validated questionnaire [26].

3.5 Daily Calcium Intake

The estimation of the daily calcium intake (DCI) was based on a frequency questionnaire [27]. Selection of items was based on the food composition diet, frequency of use, and relative importance of food items as a calcium source. The total number of foods was 30 items. The questionnaire included the following food items: milk and dairy products, including calcium-enriched items such as yogurt, cheese, and chocolate. Items such as eggs, meat, fish, cereals, bread, vegetables, and fruits were also included. Adequacy of calcium in the subjects was assessed using the adequate intake guidelines of 1000 mg of calcium.

3.6 Daily Protein Intake

The estimation of the daily protein intake (DPI) was based on a frequency questionnaire [28].

3.7 Sleep Quality

The estimation of the sleep quality index was evaluated using a self-reported questionnaire [29].

3.8 Bone Measurements

BMC (in g), BMD (in g/cm²) and TBS (L1–L4 region) were determined for each individual. BMD measurements were completed for the whole body (WB), the lumbar spine (L1–L4), the total hip (TH) and the femoral neck (FN) using the instrument described earlier (GE-Lunar, iDXA). The coefficients of variation were <1% for BMC and BMD in our laboratory [20, 23, 25].

3.9 Statistical Analysis

The means and standard deviations were calculated for all clinical data and for the bone measurements. Comparisons between the two groups were made after checking for Gaussian distribution. Comparisons between the 2 groups (overweight and normal-weight) were made after checking for Gaussian distribution. If Gaussian

distribution was found, parametric unpaired *t*-tests were used. In other cases, Mann-Whitney U-tests were used. Associations between clinical and bone data were given as Pearson's correlation coefficients (for normally distributed variables) or Spearman's correlation coefficients (for non-normally distributed variables). Multiple linear regression analysis models were used to test the relationship of DXA variables with lean mass and VO₂ max. Data were analyzed with Number Cruncher Statistical System (NCSS, 2001). A level of significance of $p < 0.05$ was used.

4 Results

4.1 Clinical Characteristics and Bone Variables of the Study Population

Among all variables, normality test failed only for physical activity level (h/week) and WB BMC (Table 1). Age, height, daily calcium intake, physical activity, VO₂ max (l/mn), PSQI and L1–L4 BMD were not significantly different between the two groups.

Table 1. Clinical characteristics and bone variables in overweight and normal-weight women

	Overweight women (n = 14)	Normal-weight women (n = 42)
Age (years)	23.6 ± 4.9	22.0 ± 2.4
Weight (kg)	78.2 ± 15.8***	54.1 ± 5.7
Height (cm)	158.9 ± 7.1	159.7 ± 5.9
BMI (kg/m ²)	30.8 ± 4.8***	21.3 ± 2.1
Lean mass (kg)	41.512 ± 6.228***	34.958 ± 3.855
Fat mass (kg)	34.525 ± 10.331***	17.165 ± 3.842
DCI (mg/d)	520 ± 173	651 ± 174
DPI (g/d)	62.6 ± 19.2	52.8 ± 19.1
Physical activity (h/week)	<i>3.7 ± 4.3</i>	<i>3.0 ± 3.0</i>
VO ₂ max (l/mn)	1.91 ± 0.38	1.75 ± 0.42
PSQI	4.2 ± 2.1	4.7 ± 2.5
WB BMC (g)	2277 ± 360*	2097 ± 282
WB BMD (g/cm ²)	1.11 ± 0.11*	1.04 ± 0.09
L1–L4 BMD (g/cm ²)	1.16 ± 0.16	1.12 ± 0.14
L1–L4 TBS	1.38 ± 0.12	1.43 ± 0.08
TH BMD (g/cm ²)	1.04 ± 0.15*	0.955 ± 0.119
FN BMD (g/cm ²)	1.02 ± 0.18*	0.935 ± 0.129

Note: Values are means ± standard deviations. In italic, Mann-Whitney U-test is used. BMI: Body Mass Index; DCI: Daily Calcium Intake; DPI: Daily Protein Intake; PA: Physical Activity; VO₂ max: Maximal Oxygen Consumption; PSQI: Pittsburgh Sleep Quality Index; WB: Whole Body; BMC: Bone Mineral Content; BMD: Bone Mineral Density; TBS: Trabecular Bone Score; TH: Total Hip; FN: Femoral Neck; *Overweight women significantly different from Normal-weight, $p < 0.05$; ***Overweight women significantly different from Normal-weight, $p < 0.001$.

Weight, BMI, lean mass, fat mass, WB BMC, WB BMD, TH BMD and FN BMD were significantly higher in overweight women compared to normal-weight women ($p < 0.05$). L1–L4 TBS was not significantly different between the two groups (Table 1).

4.2 Correlations Between Clinical Characteristics and Bone Variables

Weight, BMI, lean mass, physical activity level and WB BMC had a non-normal distribution (Table 2). Weight, height, BMI, lean mass and fat mass were positively correlated to BMC and BMD values but not to TBS values. VO_2 max (l/min) was positively correlated to bone variables (BMC and BMD) and to TBS (Table 2). Age, physical activity, DCI, DPI and PSQI were not significantly correlated to TBS.

Table 2. Correlations between clinical characteristics and bone variables

	WB BMC	WB BMD	L1–L4 BMD	TH BMD	FN BMD	TBS
Age (years)	<i>0.21</i>	<i>0.27*</i>	<i>0.35***</i>	0.12	0.14	0.03
Weight (kg)	<i>0.60***</i>	<i>0.64***</i>	<i>0.43***</i>	<i>0.59***</i>	<i>0.59***</i>	<i>0.08</i>
Height (cm)	<i>0.65***</i>	<i>0.41**</i>	<i>0.42**</i>	<i>0.33*</i>	<i>0.43***</i>	0.24
BMI (kg/m^2)	<i>0.40***</i>	<i>0.51***</i>	<i>0.28*</i>	<i>0.49***</i>	<i>0.46***</i>	<i>0.01</i>
Lean mass (kg)	<i>0.77***</i>	<i>0.71***</i>	<i>0.58***</i>	<i>0.66***</i>	<i>0.67***</i>	<i>0.15</i>
Fat mass (kg)	<i>0.41***</i>	<i>0.49***</i>	<i>0.26*</i>	<i>0.46***</i>	<i>0.45***</i>	0.08
DCI (mg/j)	<i>0.12</i>	-0.13	0.13	-0.06	-0.01	0.12
DPI (g/j)	<i>0.24</i>	0.09	<i>0.45*</i>	0.03	0.02	0.32
Physical activity (h/week)	<i>0.31*</i>	<i>0.19</i>	<i>0.46*</i>	<i>0.14</i>	<i>0.07</i>	<i>0.09</i>
VO_2 max (l/min)	<i>0.67***</i>	<i>0.68***</i>	<i>0.61***</i>	<i>0.67***</i>	<i>0.70***</i>	<i>0.36*</i>
PSQI	-0.08	0.05	-0.09	-0.01	-0.06	-0.00

Note: The results shown represent r values. In italic, spearman test is used. BMI: Body Mass Index; WB: Whole Body; BMC: Bone Mineral Content; BMD: Bone Mineral Density; TBS: Trabecular Bone Score; TH: Total Hip; FN: Femoral Neck; DCI: Daily Calcium Intake; DPI: Daily Protein Intake; PA: Physical Activity; VO_2 max: Maximal Oxygen Consumption; PSQI: Pittsburgh Sleep Quality Index; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

4.3 Correlations Between TBS and BMC/BMD Values

TBS was positively correlated to WB BMC ($r = 0.47$; $p < 0.001$), WB BMD ($r = 0.45$; $p < 0.001$), L1–L4 BMD ($r = 0.59$; $p < 0.001$), TH BMD ($r = 0.51$; $p < 0.001$) and FN BMD ($r = 0.52$; $p < 0.001$).

4.4 Multiple Linear Regression Models

Lean mass was a stronger determinant of WB BMC and WB BMD than VO₂ max (L/mn). VO₂ max (L/mn) was a stronger determinant of L1–L4 BMD, TH BMD and FN BMD than lean mass.

5 Discussion

This study conducted on a group of young Lebanese women mainly shows that being overweight is not associated with higher TBS values, and that body weight (Fig. 3) and body mass index (Fig. 1) are not positive determinants of TBS. Up to our knowledge, this is the first study to compare trabecular bone score in overweight and normal-weight young women. Moreover, this is the first study to explore the relations between many traditional bone mass determinants (weight, lean mass, BMI, daily calcium intake and physical activity) and TBS in young women. WB BMC, WB BMD, TH BMD and FN BMD were higher in overweight women compared to normal-weight women. This result is in line with those of our previous studies [18–21]. TBS values were not significantly different between the two groups. Regarding the correlations, weight, lean mass, fat mass and BMI were positively correlated to BMC and BMD values but not to TBS values. Moreover, TBS was poorly correlated to BMD values (*r* values range between 0.45 and 0.59) suggesting that BMD and TBS may reflect different bone properties. Therefore, TBS and BMD seem to not have the same positive determinants. A recent study conducted on adolescent girls with anorexia nervosa showed that TBS values were low and that TBS showed evidence of degraded microarchitecture in over 40% of the study sample [17]. Donaldson et al. [17] have also revealed that weight, lean mass, fat mass and BMI were positively correlated to TBS in adolescent girls with anorexia nervosa. Their result is in contrast with our results. Accordingly, it seems that the correlations between morphological characteristics (weight, lean mass, fat mass and BMI) and TBS values are influenced by the weight status of the studied population. Based on this, the relation between BMI and TBS is definitely not linear in women. Physical activity (h/week) was positively correlated to WB BMC and to L1–L4 BMD. VO₂ max (l/mn) was positively correlated to all bone variables including TBS. Our study confirms the results of our previous studies in which we have shown positive associations between VO₂ max (l/mn) and bone variables [23, 25]. In fact, VO₂ max (L/mn) was the best determinant of lumbar spine and hip BMD in our study. Based on our results, it seems that maximizing VO₂ max (L/mn) during growth and early adulthood may help to prevent osteopenia and osteoporosis later in life. DPI was positively correlated to L1–L4 BMD in our study. The mechanisms by which protein intake influences bone mass have been previously described [30, 31]. BMI was positively correlated L1–L4 BMD (Fig. 2). Body weight was positively correlated to L1–L4 BMD (Fig. 4). In this report, we didn't find a clear relation between sleep quality and bone variables. The mechanisms by which sleep quality influences bone variables have been previously described [32, 33]. However, further studies are needed to completely understand the relation between sleep and bone strength. Our study has several limitations. The first limitation is the cross-sectional nature of this study. The

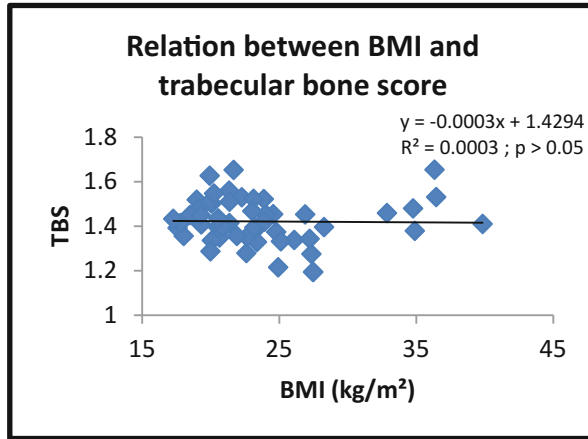


Fig. 1. Relation between body mass index and trabecular bone score

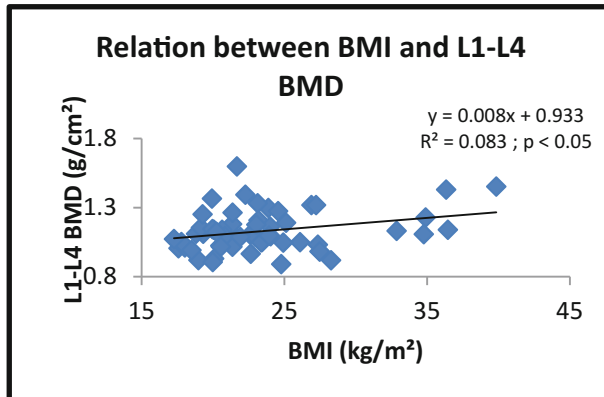


Fig. 2. Relation between body mass index and lumbar spine bone mineral density

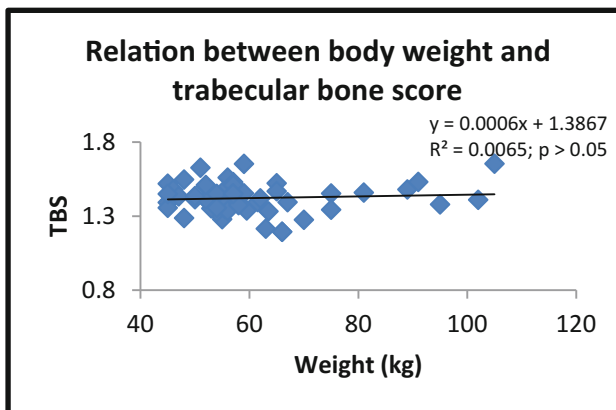


Fig. 3. Relation between body weight and trabecular bone score

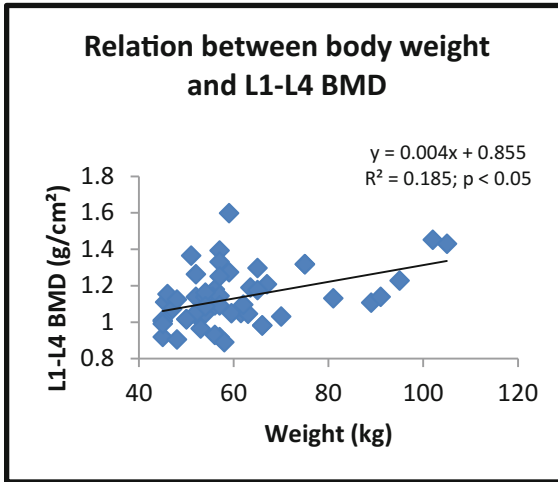


Fig. 4. Relation between body weight and lumbar spine bone mineral density

second limitation is the low number of subjects in each group (overweight and normal-weight). The third limitation is that there's a big difference between the 2 groups regarding the number of subjects (42 vs. 14 subjects). The fourth limitation is the 2-dimensional nature of DXA [4, 5]. Finally, the use of self-reported questionnaires to evaluate dietary intakes and sleep may be considered as a limitation. However, up to our knowledge, it is the first study to explore the influence of weight status on trabecular bone score in young women. Identifying the determinants of TBS in young women would be useful for the prevention of osteoporotic fractures later in life.

6 Conclusion

In conclusion, this study suggests that being overweight does not positively affect trabecular bone score values in young women. Further research is needed to identify the positive determinants of TBS in young women. In addition, further larger studies are needed to fully understand the effects of obesity and overweight on TBS in adult women and to confirm whether these effects are different from those on BMC and BMD.

Acknowledgments. This study was supported by a grant from the research council of the University of Balamand, Lebanon. Moreover, we gratefully acknowledge Lina Rahmé El Hage for her help in improving the quality of this manuscript.

Conflicts of Interest. The authors state that they have no conflicts of interest.

References

1. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis and Therapy: Osteoporosis prevention, diagnosis and therapy. *JAMA* **285**, 785–795 (2001)
2. Russo, C.R., Lauretani, F., Seeman, E., et al.: Structural adaptations to bone loss in aging men and women. *Bone* **38**(1), 112–118 (2006)
3. El Hage, R., Jacob, C., Moussa, E., et al.: Relative importance of lean mass and fat mass on bone mineral density in a group of Lebanese postmenopausal women. *J Clin. Densitom.* **14** (3), 326–331 (2011)
4. Fonseca, H., Moreira-Gonçalves, D., Coriolano, H.J., et al.: Bone quality: the determinants of bone strength and fragility. *Sports Med.* **44**(1), 37–53 (2014)
5. Bousson, V., Bergot, C., Sutter, B., et al.: Trabecular bone score (TBS): available knowledge, clinical relevance, and future prospects. *Osteoporos. Int.* **23**(5), 1489–1501 (2012)
6. Silva, B.C., Leslie, W.D., Resch, H., et al.: Trabecular bone score: a noninvasive analytical method based upon the DXA image. *J. Bone Miner. Res.* **29**(3), 518–530 (2014)
7. Iki, M., Tamaki, J., Kadowaki, E., et al.: Trabecular bone score (TBS) predicts vertebral fractures in Japanese women over 10 years independently of bone density and prevalent vertebral deformity: the Japanese population-based osteoporosis (JPOS) cohort study. *J. Bone Miner. Res.* **29**(2), 399–407 (2014)
8. Silva, B.C., Bilezikian, J.P.: Trabecular bone score: perspectives of an imaging technology coming of age. *Arq. Bras. Endocrinol. Metabol.* **58**(5), 493–503 (2014)
9. Silva, B.C., Broy, S.B., Boutroy, S., et al.: Fracture risk prediction by non-BMD DXA measures: the 2015 ISCD official positions part 2: trabecular bone score. *J. Clin. Densitom.* **18**(3), 309–330 (2015)
10. Roux, J.P., Wegrzyn, J., Boutroy, S., et al.: The predictive value of trabecular bone score (TBS) on whole lumbar vertebrae mechanics: an ex vivo study. *Osteoporos. Int.* **24**(9), 2455–2460 (2013)
11. Krueger, D., Fidler, E., Libber, J., et al.: Spine trabecular bone score subsequent to bone mineral density improves fracture discrimination in women. *J. Clin. Densitom.* **17**(1), 60–65 (2014)
12. Harvey, N.C., Glüer, C.C., Binkley, N., et al.: Trabecular bone score (TBS) as a new complementary approach for osteoporosis evaluation in clinical practice. *Bone* **78**, 216–224 (2015)
13. Leslie, W.D., Krieg, M.A., Hans, D., Manitoba Bone Density Program: Clinical factors associated with trabecular bone score. *J. Clin. Densitom.* **16**(3), 374–379 (2013)
14. Leslie, W.D., Aubry-Rozier, B., Lix, L.M., et al.: Spine bone texture assessed by trabecular bone score (TBS) predicts osteoporotic fractures in men: the Manitoba Bone Density Program. *Bone* **67**, 10–14 (2014)
15. Leslie, W.D., Johansson, H., Kanis, J.A., et al.: Lumbar spine texture enhances 10-year fracture probability assessment. *Osteoporos. Int.* **25**(9), 2271–2277 (2014)
16. Hans, D., Goertzen, A., Krieg, M.A., et al.: Bone microarchitecture assessed by TBS predicts hip, clinical spine and all osteoporotic fractures independently of BMD in 22234 women aged 50 and older: the Manitoba prospective study. *J. Bone Miner. Res.* **26**(11), 2762–2769 (2011)
17. Donaldson, A.A., Feldman, H.A., O'Donnell, J.M., et al.: Spinal bone texture assessed by trabecular bone score in adolescent girls with anorexia nervosa. *J. Clin. Endocrinol. Metab.* **100**(9), 3436–3442 (2015)

18. El Hage, R., Bachour, F., Sebaaly, A., et al.: The influence of weight status on radial bone mineral density in Lebanese women. *Calcif. Tissue Int.* **94**(4), 465–467 (2014)
19. El Hage, R., Bachour, F., Khairallah, W., et al.: Lumbar spine bone mineral density in obese, overweight, and normal-weight Lebanese postmenopausal women. *J. Clin. Densitom.* **17**(1), 215–216 (2014)
20. El Hage, R., El Hage, Z., Moussa, E., et al.: Geometric indices of hip bone strength in obese, overweight, and normal-weight adolescent girls. *J. Clin. Densitom.* **16**(3), 313–319 (2013)
21. El Hage, R., Bachour, F., Khairallah, W., et al.: The influence of obesity and overweight on hip bone mineral density in Lebanese women. *J. Clin. Densitom.* **17**(1), 216–217 (2014)
22. El Hage, R., Khairallah, W., Bachour, F., et al.: Influence of age, morphological characteristics, and lumbar spine bone mineral density on lumbar spine trabecular bone score in Lebanese women. *J. Clin. Densitom.* **17**(3), 434–435 (2014)
23. Zakhem, E., Ayoub, M.L., Zunquin, G., et al.: Physical performance and trabecular bone score in a group of young lebanese women. *J. Clin. Densitom.* **18**(2), 271–272 (2015)
24. Cole, T.J., Bellizzi, M.C., Flegal, K.M., et al.: Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**(7244), 1240–1243 (2000)
25. El Hage, R., Zakhem, E., Theunynck, D., et al.: Maximal oxygen consumption and bone mineral density in a group of young Lebanese adults. *J. Clin. Densitom.* **17**(2), 320–324 (2014)
26. Deheeger, M., Rolland-Cachera, M.F., Fontvieille, A.M.: Physical activity and body composition in 10 year old French children: linkages with nutritional intake? *Int. J. Obes. Relat. Metab. Disord.* **21**, 372–379 (1997)
27. Fardellone, P., Sebert, J.L., Bouraya, M., et al.: Evaluation of the calcium content of diet by frequential self-questionnaire. *Rev. Rhum. Mal. Osteoartic.* **58**, 99–103 (1991)
28. Morin, P., Herrmann, F., Ammann, P., et al.: A rapid self-administered food frequency questionnaire for the evaluation of dietary protein intake. *Clin. Nutr.* **24**, 768–774 (2005)
29. Buysse, D.J., Reynolds 3rd, C.F., Monk, T.H., et al.: The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* **28**, 193–213 (1989)
30. Bonjour, J.P., Chevalley, T., Ferrari, S., et al.: The importance and relevance of peak bone mass in the prevalence of osteoporosis. *Salud Publica Mex.* **51**(1 Suppl.), S5–S17 (2009)
31. Bonjour, J.P., Chevalley, T., Rizzoli, R., et al.: Gene-environment interactions in the skeletal response to nutrition and exercise during growth. *Med. Sport Sci.* **51**, 64–80 (2007)
32. Zakhem, E., El Hage, R., Zunquin, G., et al.: Sleep quality is a determinant of hip bone mineral density in a group of young Lebanese men. *J. Med. Liban.* **62**(4), 213–216 (2014)
33. Albayrak, I., Aydogmus, M., Ozerbil, O.M., et al.: The association between bone mineral density, quality of life, quality of sleep and fatigue. *Acta Clin. Belg.* **71**(2), 92–98 (2016). <https://doi.org/10.1179/2295333715Y.0000000061>



Sarcopenia and Hip Structure Analysis Variables in a Group of Lebanese Postmenopausal Women

Riad Nasr^{1,2}, Eric Watelain^{3,4}, Antonio Pinti⁵ (✉), Hayman Saddik¹, Ghassan Maalouf⁶, Abdel-Jalil Berro^{1,7}, Abir Alwan^{1,8}, César El Khoury^{1,5}, Ibrahim Fayad¹, and Rawad El Hage¹

¹ Department of Physical Education, University of Balamand, El-Koura, Lebanon

² DeVisu - Design, Visuel, Urbain, EA 2445, UVHC, Valenciennes, France

³ CNRS, UMR 8201, Valenciennes, France

⁴ LAMIH-Département SHV, Université Lille Nord de France, UVHC, Valenciennes, France

⁵ I3MTO, EA4708, Université d'Orléans, Orléans, France
antonio.pinti@univ-valenciennes.fr

⁶ Faculty of Medicine, Bellevue University Medical Center, Saint Joseph University, Mansourieh, Lebanon

⁷ EA-3300, APERE, University of Picardie Jules Verne, Amiens, France

⁸ Laboratoire Mouvement Sport Santé, UFR APS, Rennes, France

Abstract. The aim of the present study was to compare hip structural analysis variables in postmenopausal women with sarcopenia and postmenopausal women with normal skeletal muscle mass index. This study included 8 postmenopausal women (aged between 65 and 84 years) with sarcopenia and 60 age-matched controls (with normal skeletal muscle mass index (SMI)). Body composition and bone mineral density (BMD) were assessed by dual-energy X-ray absorptiometry (DXA). Weight, lean mass, Body mass index, femoral neck cross-sectional area (FN CSA), FN section modulus (Z), FN cross sectional moment of inertia (CSMI), intertrochanteric (IT) CSA, IT Z, IT CSMI, IT cortical thickness (CT), femoral shaft (FS) CSA, FS Z and FS CSMI were significantly higher ($p < 0.05$) in women with normal SMI compared to women with sarcopenia. In the whole population, SMI was positively correlated to IT CSA, IT Z, IT CSMI, IT CT, FS CSA, FS Z, FS CSMI, FS CT but negatively correlated to IT buckling ratio (BR) and FS BR. The present suggests that sarcopenia negatively affects hip bone strength indices in postmenopausal women.

Keywords: Sarcopenia · DXA imaging · Strength indices · Fracture risk

1 Introduction

Muscle mass decreases with age, leading to “sarcopenia” or low muscle mass, in elderly people [1, 2]. With the increasing of life span over the past several decades, this phenomenon is becoming more remarkable and thus, has a more significant impact on

society [3]. Lean mass is a strong determinant of bone mineral density (BMD) and geometric indices of hip bone strength in the elderly [4–9]. Recently, sarcopenia has been related to low BMD values in the elderly [10, 11]. However, bone strength is not only influenced by BMD but is also influenced by bone geometry [12–14]. Several studies have shown that hip structure analysis (HSA) variables can predict incident hip fracture risk in postmenopausal women [15–23]. We have recently shown that sarcopenia negatively affects hip bone strength indices in elderly men [24]. However, the relation between sarcopenia and hip bone strength indices in elderly women remains unclear. The aim of the present study was to compare HSA variables in postmenopausal women with sarcopenia and postmenopausal women with normal skeletal muscle mass index.

2 Methods

This study included 8 postmenopausal women (aged between 65 and 84 years) with sarcopenia and 60 age-matched controls (with normal skeletal muscle mass index). Body composition and bone mineral density (BMD) were assessed by dual-energy X-ray absorptiometry (Hologic QDR-4500 W, Waltham, MA USA). We used the skeletal muscle mass index (appendicular skeletal mass (ASM)/height²) to define sarcopenia as previously described [1]. A skeletal muscle mass index (SMI) < 5.5 kg/m² was defined as the cut point for sarcopenia [1]. To evaluate hip bone geometry, DXA scans were analyzed at the femoral neck (FN), the intertrochanteric region (IT), and the femoral shaft (FS) by the Hip Structure Analysis (HSA) program (Fig. 1). Cross sectional area (CSA), an index of axial compression strength, section modulus (Z), an index of bending strength, cross sectional moment of inertia (CSMI), an index of structural rigidity, cortical thickness (CT) and buckling ratio (BR), an index of bone geometric instability, were measured from bone mass profiles [4, 12]. The means and standard deviations were calculated for all clinical data and for the bone measurements. Comparisons between the two groups (sarcopenic women and women with normal SMI) were made after checking for Gaussian distribution. If Gaussian distribution was found, parametric unpaired t-tests were used. In other cases, Mann-Whitney U tests were used. Associations between clinical characteristics and bone variables were given as Pearson correlation coefficients. HSA variables were compared between the two groups (sarcopenic women and women with normal SMI) after adjustment for lean mass using a one-way analysis of covariance (ANCOVA). Data were analyzed with Number Cruncher Statistical System (NCSS, 2001). A level of significance of $p < 0.05$ was used.

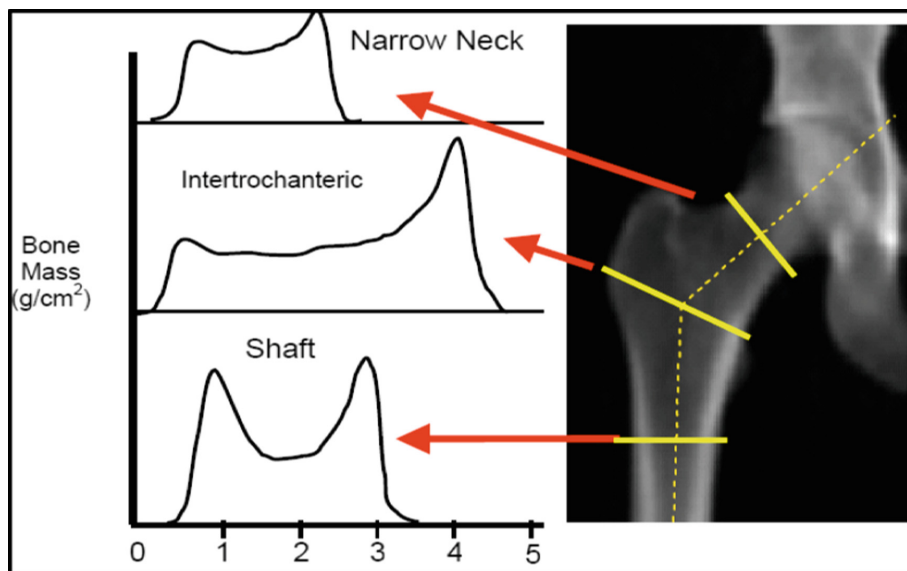


Fig. 1. Hip image from a Hologic DXA scanner showing positions of thin analysis regions across the femur at the neck (NN region), intertrochanteric (IT) and shaft. On the left are shown typical bone mass profiles used in measurements of geometric properties.

3 Results

Weight, Body Mass Index (BMI), lean mass (LM), Fat Mass (FM), Femoral Neck (FN) CSA, FN Z, FN CSMI, IT CSA, IT Z, IT CSMI, IT CT, FS CSA, FS Z and FS CSMI were significantly higher ($p < 0.05$) in women with normal SMI compared to women with sarcopenia (Table 1). IT BR was significantly higher ($p < 0.05$) in women with sarcopenia compared to women with normal SMI.

In the whole population (Table 2), SMI was positively correlated to IT CSA ($r = 0.30$; $p < 0.05$), IT Z ($r = 0.28$; $p < 0.05$), IT CSMI ($r = 0.24$; $p < 0.05$), IT CT ($r = 0.29$; $p < 0.05$), FS CSA ($r = 0.36$; $p < 0.01$), FS Z ($r = 0.36$; $p < 0.01$), FS CSMI ($r = 0.32$; $p < 0.01$), FS CT ($r = 0.32$; $p < 0.01$) but negatively correlated to IT BR ($r = -0.31$; $p < 0.05$) and FS BR ($r = -0.23$; $p < 0.05$). The buckling ratio is an estimate of relative cortical thickness (subperiosteal radius/cortical thickness). A higher Buckling Ratio (BR) value means greater instability and is associated with fracture risk [25–28].

After adjusting for lean mass using an ANCOVA, there were no significant differences between the two groups regarding bone variables (WB BMC, WBBMD, L1-L4 BMD, TH BMD, FN BMD, NN CSA, NN Z, IT CSA, IT Z, FS CSA and FS Z).

Table 1. Clinical characteristics and bone variables in sarcopenic women and normal women

	Women with normal SMI (n = 60)	Sarcopenic women (n = 8)	P-value
Age (years)	71.633 ± 4.780	75.375 ± 6.610	0.051
Weight (kg)	70.233 ± 12.500**	55.875 ± 9.141	0.003
Height (m)	1.519 ± 0.0596	1.504 ± 0.0752	0.503
BMI (kg/m ²)	30.343 ± 4.745**	24.653 ± 3.059	0.002
FM (kg)	27.520 ± 7.810*	20.578 ± 6.490	0.019
LM (kg)	39.986 ± 5.344***	31.236 ± 2.827	<0.001
WB BMD (g/cm ²)	0.884 ± 0.0841	0.834 ± 0.0618	0.106
WB BMC (kg)	1.528 ± 0.280*	1.285 ± 0.202	0.021
L1–L4 BMD (g/cm ²)	0.783 ± 0.158	0.746 ± 0.0821	0.827
Total hip BMD (g/cm ²)	0.727 ± 0.124*	0.616 ± 0.0894	0.017
Femoral neck BMD (g/cm ²)	0.632 ± 0.100**	0.532 ± 0.0871	0.009
NN CSA (cm ²)	2.214 ± 0.405**	1.821 ± 0.212	0.009
NN Z (cm ³)	1.072 ± 0.284*	0.801 ± 0.159	0.010
IT CSA (cm ²)	3.726 ± 0.836**	2.882 ± 0.356	0.007
IT Z (cm ³)	3.090 ± 0.879**	2.208 ± 0.399	0.003
IT BR	11.1 ± 2.7*	13.1 ± 2.6	0.04
FS CSA (cm ²)	3.520 ± 0.647**	2.853 ± 0.343	0.006
FS Z (cm ³)	2.094 ± 0.419***	1.516 ± 0.237	<0.001

WB BMC: Whole Body Bone Mineral Content; WB BMD: Whole Body Bone Mineral Density; NN CSA: Narrow Neck Cross-Sectional Area; Z: Section modulus; IT: Intertrochanteric; FS: Femoral shaft; BR: Buckling Ratio; *p < 0.05; **p < 0.01; *** p < 0.001.

Table 2. Correlations between SMI and bone variables

	SMI (kg/m ²)
WB BMC (g)	0.29*
WB BMD (g/cm ²)	0.26*
L1–L4 BMD (g/cm ²)	0.08
TH BMD (g/cm ²)	0.32**
FN BMD (g/cm ²)	0.25*
NN CSA (cm ²)	0.21
NN Z (cm ³)	0.18
IT CSA (cm ²)	0.30*
IT Z (cm ³)	0.28*
IT BR	-0.31*
FS CSA (cm ²)	0.36**
FS Z (cm ³)	0.36**
FS BR	-0.23*

FN: Femoral Neck; TH: Total Hip; WB BMC: Whole Body Bone Mineral Content; WB BMD: Whole Body Bone Mineral Density; NN CSA: Narrow Neck Cross-Sectional Area; Z: Section modulus; BR: Buckling Ratio; IT: Intertrochanteric; FS: Femoral Shaft; *p < 0.05; **p < 0.01.

4 Conclusion

The present suggests that sarcopenia negatively affects hip bone strength indices in postmenopausal women. Implementing strategies to increase SMI in postmenopausal women may be useful for preventing osteoporotic fractures.

Conflicts of Interest. The authors state that they have no conflicts of interest.

References

1. Baumgartner, R.N., Koehler, K.M., Gallagher, D., et al.: Epidemiology of sarcopenia among the elderly in new Mexico. *Am. J. Epidemiol.* **147**(8), 755–763 (1998)
2. Cruz-Jentoft, A.J., Baeyens, J.P., Bauer, J.M., et al.: Sarcopenia: European consensus on definition and diagnosis: report of the european working group on sarcopenia in older people. *Age Ageing* **39**(4), 412–423 (2010)
3. Morley, J.E., Abbatecola, A.M., Argiles, J.M., et al.: Sarcopenia with limited mobility: an international consensus. *J. Am. Med. Dir. Assoc.* **12**(6), 403–409 (2011)
4. El Hage, R., Theunynck, D., Rocher, E., et al.: Geometric indices of hip bone strength in overweight and control elderly men. *J. Med. Liban.* **62**(3), 150–155 (2014)
5. El Hage, R., Mina, F., Ayoub, M.L., et al.: Relative importance of lean mass and fat mass on bone mineral density in a group of Lebanese elderly men. *J. Med. Liban.* **60**(3), 136–141 (2012)
6. Taaffe, D.R., Villa, M.L., Holloway, L., et al.: Bone mineral density in older non-Hispanic Caucasian and Mexican-American women: relationship to lean and fat mass. *Ann. Hum. Biol.* **27**(4), 331–344 (2000)
7. Chiu, G.R., Araujo, A.B., Travison, T.G., et al.: Relative contributions of multiple determinants to bone mineral density in men. *Osteoporos. Int.* **20**(12), 2035–2047 (2009)
8. Beck, T.J., Petit, M.A., Wu, G., et al.: Does obesity really make the femur stronger? BMD, geometry, and fracture incidence in the women’s health initiative-observational study. *J. Bone Miner. Res.* **24**(8), 1369–1379 (2009)
9. Travison, T.G., Araujo, A.B., Esche, G.R., et al.: Lean mass and not fat mass is associated with male proximal femur strength. *J. Bone Miner. Res.* **23**(2), 189–198 (2008)
10. Edwards, M.H., Dennison, E.M., Aihie Sayer, A., et al.: Osteoporosis and sarcopenia in older age. *Bone* **80**, 126–130 (2015)
11. Pereira, F.B., Leite, A.F., de Paula, A.P.: Relationship between pre-sarcopenia, sarcopenia and bone mineral density in elderly men. *Arch. Endocrinol. Metab.* **59**(1), 59–65 (2015)
12. Beck, T.J., Ruff, C.B., Warden, K.E., et al.: Predicting femoral neck strength from bone mineral data. *Struct. Approach Invest Radiol.* **25**(1), 6–18 (1990)
13. Beck, T.J., Broy, S.B.: Measurement of hip geometry-technical background. *J. Clin. Densitom* **18**(3), 331–337 (2015)
14. Bouxsein, M.L., Karasik, D.: Bone geometry and skeletal fragility. *Curr. Osteoporos. Rep.* **4**(2), 49–56 (2006)
15. LaCroix, A.Z., Beck, T.J., Cauley, J.A., et al.: Hip structural geometry and incidence of hip fracture in postmenopausal women: what does it add to conventional bone mineral density? *Osteoporos. Int.* **21**(6), 919–929 (2010)
16. Reider, L., Beck, T.J., Hochberg, M.C., et al.: Women with hip fracture experience greater loss of geometric strength in the contralateral hip during the year following fracture than age-matched controls. *Osteoporos. Int.* **21**(5), 741–750 (2010)

17. Kaptoge, S., Beck, T.J., Reeve, J., et al.: Prediction of incident hip fracture risk by femur geometry variables measured by hip structural analysis in the study of osteoporotic fractures. *J. Bone Miner. Res.* **23**(12), 1892–1904 (2008)
18. Melton III, L.J., Beck, T.J., Amin, S., et al.: Contributions of bone density and structure to fracture risk assessment in men and women. *Osteoporos. Int.* **16**(5), 460–467 (2005)
19. Szulc, P., Duboeuf, F., Schott, A.M., et al.: Structural determinants of hip fracture in elderly women: re-analysis of the data from the EPIDOS study. *Osteoporos. Int.* **17**(2), 231–236 (2006)
20. Leslie, W.D., Pahlavan, P.S., Tsang, J.F., et al.: Prediction of hip and other osteoporotic fractures from hip geometry in a large clinical cohort. *Osteoporos. Int.* **20**(10), 1767–1774 (2009)
21. El-Kaissi, S., Pasco, J.A., Henry, M.J., et al.: Femoral neck geometry and hip fracture risk: the Geelong osteoporosis study. *Osteoporos. Int.* **16**(10), 1299–1303 (2005)
22. Crabtree, N.J., Kroger, H., Martin, A., et al.: Improving risk assessment: hip geometry, bone mineral distribution and bone strength in hip fracture cases and controls. The EPOS study. European prospective osteoporosis study. *Osteoporos. Int.* **13**(1), 48–54 (2002)
23. Ahlborg, H.G., Nguyen, N.D., Nguyen, T.V., et al.: Contribution of hip strength indices to hip fracture risk in elderly men and women. *J. Bone Miner. Res.* **20**(10), 1820–1827 (2005)
24. Nasr, R., Watelain, E., Pinti, A., et al.: Sarcopenia and hip structure analysis variables in a group of elderly men. *J. Clin. Densitom.* (2016, in press)
25. Anitha, D., Lee, T.: Assessing bone quality in terms of bone mineral density, buckling ratio and critical fracture load. *J. Bone. Metab.* **21**(4), 243–247 (2014)
26. Sheu, Y., Cauley, J.A., Patrick, A.L., et al.: Risk factors for fracture in middle-age and older-age men of African descent. *J. Bone Miner. Res.* **29**(1), 234–241 (2014)
27. Duan, Y., Beck, T.J., Wang, X.F., et al.: Structural and biomechanical basis of sexual dimorphism in femoral neck fragility has its origins in growth and aging. *J. Bone Miner. Res.* **18**(10), 1766–1774 (2003)
28. Kaptoge, S., Dalzell, N., Loveridge, N., et al.: Effects of gender, anthropometric variables, and aging on the evolution of hip strength in men and women aged over 65. *Bone* **32**(5), 561–570 (2003)



Feet Fidgeting Detection Based on Accelerometers Using Decision Tree Learning and Gradient Boosting

Julien Esseiva¹(✉), Maurizio Caon¹, Elena Mugellini¹,
Omar Abou Khaled¹, and Kamiar Aminian²

¹ University of Applied Sciences and Arts Western Switzerland,
Fribourg, Switzerland
{julien.esseiva, maurizio.caon, elena.mugellini,
omar.aboukhaled}@hefr.ch

² Laboratory of Movement Analysis and Measurement,
Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland
kamiar.aminian@epfl.ch

Abstract. Detection of fidgeting activities is a field which has not been much explored as of now. Studies have shown that fidgeting has a beneficial impact on people's healthiness as it burns a significant amount of energy. Being able to detect when someone is fidgeting would allow to study more closely the health impact of fidgeting. The purpose of this work is to propose an algorithm being able to detect feet fidgeting period of subjects while sitting using 3-D accelerometers on both shoes. Initial results on data from 5 subjects collected during this work shows an accuracy of 95% for a classification between sitting with fidgeting and sitting without fidgeting.

Keywords: Fidgeting detection · Decision tree · Boosting · Accelerometers
Footwear · Wearable · Machine learning

1 Introduction

Physical inactivity is more and more frequent in today society. It is known that sedentary has a negative impact on health, notably weight increase as energy is conserved as body fat. Health specialists therefore recommend to people to do physical activities to stay healthy. Recent studies have shown that people doing a light physical activity like walking, household activities or fidgeting could add up to several hundred of calories burnt per day, which has an impact on human well-being. While the field of activity and gesture recognition has seen a lot of research the past years, mainly due to the growing of fields such as human-computer interaction (HCI) or health-monitoring devices, the problem of detecting fidgeting periods has seen much less research. Activity and gesture recognition might initially seem to be a problem similar to fidgeting detection. There are however some key differences, mainly in the nature of movements. Activities are spanning on a longer period while fidgeting is generally short and spontaneous. This has an impact on the discriminative power of features, the

mean of a signal is frequently used in posture recognition as it averages out over a long time, while it might not be as important for fidgeting detection due to the short and spontaneous nature of the movement. An algorithm allowing to detect whether a subject is fidgeting is of great interest, it would allow to further assess the impact of fidgeting on energy consumption or it could be used to encourage fidgeting. Therefore, the purpose of this work is to propose an algorithm detecting periods where the subject is fidgeting with its feet by using smart shoes composed of tri-axial accelerometers. First a selection of similar work is presented. The dataset as well as the data collection protocol and instrumented shoe system is discussed. The algorithms chapter presents the pipeline of the data, the features extracted from the signals as well as trained models. Tests done during the work, which include determining an optimal window size and reproduction of a baseline algorithm are then presented, followed by a discussion of the result. The work ends with a conclusion and possible outlook.

2 Related Works

Various algorithms were proposed for the tasks of activity and gesture recognition.

Ayumi explored the application of Gradient boosting to action recognition over multiple dataset using a Kinect and depth sensor camera. While the nature of the data is different in our work, Ayumi compared gradient boosting to methods such as SVM or naive Bayes. The gradient boosting method outperformed the other two in most cases [3]. Zhang et al. worked on activity classification using decision tree and tri-axial accelerometers on instrumented shoe sensors [2]. El Achkar et al. proposed a system for activity classification using Instrumented shoe and a tree-based algorithm [5]. In these two studies data are similar to our work while movements are different, which impacts the choice of features and windowing. Tapia presents a detailed analysis of activity recognition using machine learning algorithm based on accelerometers data, exploring questions such as the processing and windowing of the signals, impact of multiple features, various models such as decision tree, naive Bayes, nearest neighbor and the differences between subject-dependent and subject-independent training [6]. Lugade et al. similarly proposed an algorithm which detects movements with a sensitivity greater than 85% using accelerometers which were placed on the waist and thigh. Amongst the activities, they were trying to detect quiet sitting and sitting with fidgeting. While the proposed algorithm could get good performances on activity recognition, they do mention that their lowest accuracy was for activities where the subject was fidgeting while sitting or standing [1]. As suggested by Lugade et al., it is necessary to do more research on human motion while fidgeting which is the focus of this work. Zhang et al. proposed an algorithm to detect feet fidgeting while the subject is sitting by using a 3D accelerometer. This is, to our knowledge, the only algorithm which is performing feet fidgeting recognition while seated and will therefore be used as a baseline algorithm for this work [4].

3 Datasets

Data in this study data were gathered using instrumented shoe sensors which were provided by [6].

3.1 Instrumented Shoe

The instrumented shoe system is composed of insoles with force sensors, a Physilog which includes 3D accelerometer, 3D gyroscope and 3D magnetometer with a sampling frequency of 200 Hz. The system can easily be inserted in any shoe. The orientation and position of Physilog sensors is not primordial as calibration of the signal is done. Figure 1 shows the instrumented shoe once set up as well as the orientation of accelerometer and gyroscope axes.



Fig. 1. Instrumented shoe setup with axes orientation

3.2 Collected Dataset

A dataset was collected for this study using the instrumented shoe in [6]. Data from 6 different subjects were collected. Each subject had to follow a predetermined protocol after a calibration of the sensors.

Collection protocol. Table 1 shows the protocol followed by each subject. They were sitting on a chair for the whole time and asked to perform four different leg gestures generally considered as fidgeting, with 5s of quiet standing in-between each fidget. During the quiet sitting period, the subject was asked not to move. This protocol was performed twice for each subject, one time for each foot. A supervisor was present during the collection protocol to note the time at which subjects were changing between tasks so that it is possible to label the data later.

The total time of the protocol is 1 min of data per subject, repeated once for each foot. Considering 6 subjects, this results in a dataset which contains a total of 12 min of data.

Table 1. Data collection protocol

Task	Duration (s)
Quiet sitting	5
Fidget 1	10
Quiet sitting	5
Fidget 2	10
Quiet sitting	5
Fidget 3	10
Quiet sitting	5
Fidget 4	10

- Fidget 1 (upper leg swinging): The subject is moving its thighs left and right, either one or both at the same time. The two feet are constantly touching the ground. This gesture could be identical for each foot, depending on the subject.
- Fidget 2 (up and down leg bouncing): The subject has his two feet on the floor and is repetitively moving the heel of one leg up and down (along the z axis) with the toe still touching the ground.
- Fidget 3 (Lower leg swinging): The heel of one leg stands on the knee of the other one. The subject was asked to move its foot.

Fidget 4 (Foot jiggling): The legs of the subject are crossed with its thighs being one over the other. The lower leg of the subject is swinging, usually along the y axis (Fig. 2).

**Fig. 2.** Illustration of the 4 types of fidgeting recorded

Collected Features. Given the multiple sensors present on the instrumented shoe, it was possible to collect a total of 15 signals per foot as listed below.

- 3D acceleration [g]
- 3D gyroscope [deg/s]
- Force of the eight insole sensors [V]
- Pressure [KPa]

Collected data were labeled to train supervised models. In that regard, signals were shown on a plot where the supervisor could specify fidgeting periods as noted during the collection protocol.

4 Algorithms

See Fig. 3.

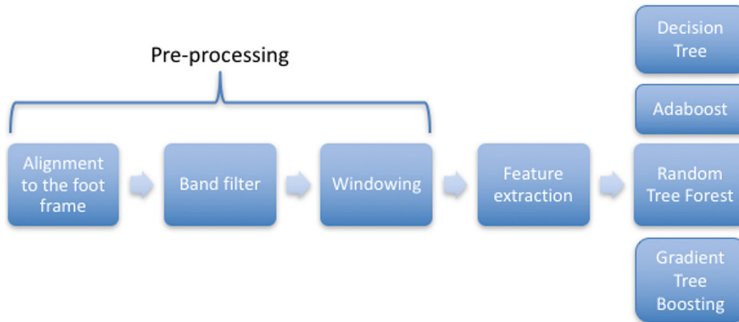


Fig. 3. Algorithm pipeline of the processed data.

4.1 Signal Processing

Filters. Raw signals extracted from sensors are usually noisy. It is important to pre-process them before extracting features from them. Removing frequencies considered to be noise is an important first step. A Butterworth bandpass filter with the following specification was applied to signals:

- Passband: [0.1 20] Hz
- Stopband: [0.01, 99] Hz
- Maximum passband ripple: 1 dB
- Minimum stopband attenuation: 60 dB

Windowing. A second important step is to split signals in multiple windows. An optimal window length is not universally defined and seems to depend on the problem to be solved. In the context of this work it is necessary to detect fidgeting which are usually short spontaneous movements. A small window therefore seems to be more appropriate to the problem. Both non-overlapping and 50% overlapping windows will be extracted from the signals to compare the two versions in terms of performance metrics.

4.2 Feature Extraction

The present work focused on using accelerometers, therefore only the six accelerometers signals, x, y and z for each foot, were used. For each signal, multiple features inspired by the work of Tapia [6] were computed as presented below. Both time and frequency features were extracted. For the frequency domain features, the 0 Hz term of the FFT was omitted when computing features related to the frequency domain. Features are extracted on a window-basis. A total of 62 features are extracted for each window and the whole set of features was used to train classification models.

FFT Peaks. The frequency component with the highest magnitude. A tuple which contains the frequency and its magnitude is extracted.

Main frequency energy ratio. This feature is computed using the Eq. 1, where mag_{hf}^2 is the highest magnitude across all frequencies of the signal. The maximum frequency is 20 Hz.

$$R = \frac{mag_{hf}^2}{\sum_{f=0}^{\max_freq} mag_f^2} \quad (1)$$

Entropy. The entropy is a commonly used feature in signal processing measuring the signal complexity. Equations 2 and 3 show how the entropy is computed, with B being the number of FFT bins.

$$H = - \sum_{b=1}^B p_b * \log_2 p_b \quad (2)$$

$$p_b = \frac{mag_b}{\sum_{j=0}^B mag_j} \quad (3)$$

Root mean square. The root mean square is a commonly used feature in statistics and is easy to compute. The signal is detrended before computing this feature, with N being the number of samples in the window and a_n is the measured acceleration at index n in the window.

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N |a_n|^2} \quad (4)$$

Absolute area. This feature is the sum of the absolute value of each sample in the window.

Absolute mean. Like the absolute area, the absolute mean corresponds to the mean of the samples in the window.

Autocovariance. The autocovariance corresponds to the cross-covariance of a signal with itself. Only the highest and the lowest values were kept for each signal which results in 2 features per signal.

Approximate entropy. The approximate entropy is a feature which is useful for capturing signal complexity and the evolution of complexity and predictability of the signal. A value close to 0 suggests that the signal is predictable and regular. The parameters $m = 2$ and $r = 0.01$ were used to compute this feature [8].

Total SVM. The signal vector magnitude of the signal is extracted as a feature by using the Eq. 5

$$SVM = \frac{1}{N} \sum_{n=1}^N \sqrt{\sum_{i=1}^{nb_axes} a_{n_i}^2} \quad (5)$$

Total absolute area. This feature corresponds to the sum of the absolute area feature of the six accelerometer signals used. It is computed once per window instead of once per signal.

4.3 Models

Multiple supervised models were trained and compared in this work, which includes decision tree, random forest, AdaBoost and gradient tree boosting. The training was done using K-fold cross-validation with $K = 10$. A grid search was performed on each model to optimize hyperparameters. Both binary and multiclass models were trained using a single subject-independent dataset. The purpose of binary models was to classify between fidgeting and no fidgeting while multiclass models should recognize each fidgeting gesture independently. For each category (binary/multiclass) of models, the four machine learning algorithms were trained once using non-overlapping window and a second time with 50% overlapping window. The choice of the optimal window size is detailed in Sect. 5.2.

5 Tests

5.1 Baseline Algorithm

Zhang et al. proposed an algorithm to detect whether a subject is sitting or sitting and fidgeting [4]. In a 2-second window, the sum of the square of the three accelerometer signals is computed for each sample in the window. Only the maximum of these values in the window is kept and signals are not preprocessed in this algorithm. If the maximum value is above a given threshold, the window is classified as fidgeting. One limitation of this algorithm is that it can't be used for multiclass classification. This algorithm was reproduced and used as a baseline comparison for the proposed algorithm. As only few information on how to compute the decision threshold was

available, a decision tree with the entropy splitting criterion and a maximum depth of 1 node was used to determine this threshold.

5.2 Optimal Window Length

Tapia [6] proposed a procedure to determine an adequate window length for a given dataset. Two features highly impacted by the window length, the Pearson correlation coefficient and FFT peaks were computed for each signal, using windows size ranging from 1s to 7s, with a step size of 0.2s. A CART Decision Tree was then trained on the two features previously mentioned for each combination of window length, window overlap and classification type. Performances were assessed using 10-fold cross-validation. The Fig. 4 shows the trend for the binary, non-overlapping windows model, other models are not shown here but similar results were obtained. The accuracy decrease as the window length increase which can be explained by the fact that increasing the window size drastically reduces the number of samples of the already small dataset therefore increasing the generalization error of the model.

Accuracy on a decision tree classifier for multiple windows length

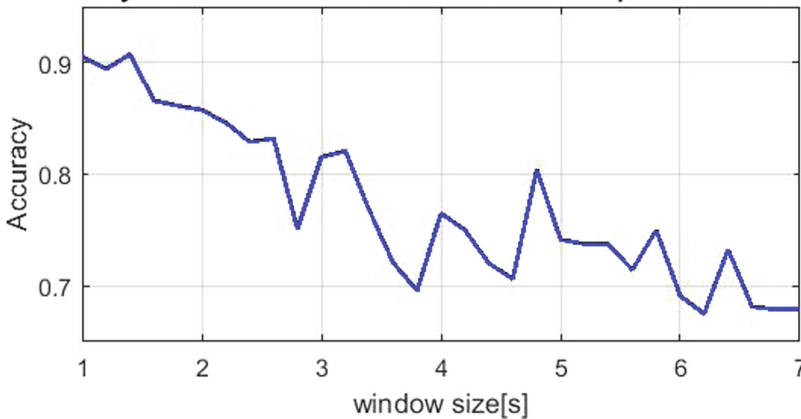


Fig. 4. Accuracy as function of window length for the binary, non-overlapping windows model.

6 Results and Discussion

Tables 2 and 3 summarize the results obtained with the best performing model resulting of the grid search using 10-fold cross-validation. Metrics are computed using a macro-average. We see that models trained on overlapping window perform slightly better overall. Performances are good for both binary and multiclass models.

Regarding the binary problem, DT perform slightly worse than other models. The Matthews correlation coefficient (MCC) is significantly lower for the baseline algorithm. Because the classes are not perfectly balanced, metrics such as accuracy are slightly skewed while the MCC is more representative of the true performances.

Table 2. Performances of binary models using rectangular windows. “No” stands for non-overlapping windows, “O” for 50% overlapping windows, “DT” is decision tree, “RF” is random forest, “AB” is AdaBoost, “GB” is gradient tree boosting.

	Accuracy		Precision		Recall		F-score		MCC	
	No	O	No	O	No	O	No	O	No	O
Base	0.80		0.83		0.81		0.81		0.58	
DT	0.89	0.93	0.89	0.93	0.89	0.93	0.89	0.93	0.72	0.83
RF	0.93	0.95	0.93	0.95	0.93	0.95	0.93	0.95	0.83	0.89
AB	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.86	0.85
GB	0.94	0.94	0.94	0.95	0.94	0.94	0.94	0.95	0.86	0.87

Table 3. Performances of the multiclass models using rectangular window.

	Accuracy		Precision		Recall		F-score	
	No	O	No	O	No	O	No	O
DT	0.77	0.82	0.77	0.82	0.77	0.82	0.77	0.82
RF	0.85	0.90	0.85	0.89	0.85	0.89	0.85	0.89
AB	0.40	0.64	0.42	0.65	0.41	0.64	0.40	0.64
GB	0.88	0.91	0.88	0.91	0.88	0.91	0.88	0.91

The multiclass problem still yields decent performances. Gradient tree boosting has the best performances overall. While models might be overfitted, it is to be noted that boosting method such as gradient tree bosting trains multiple shallow trees and is therefore resilient to overfitting.

Figure 5 shows the confusion matrix for the GB model with 50% overlapping window of length 1.2s, which results in a dataset of approximately 1200 samples. While other matrices are not presented here, the error distribution is similar in each one of them. We see that there is a confusion between the gesture “Lower leg swinging”

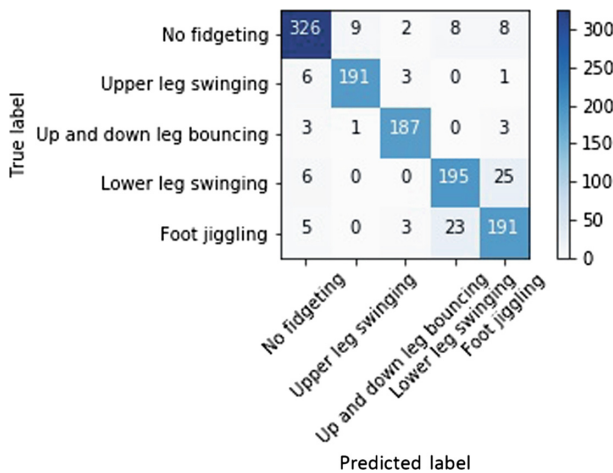


Fig. 5. Confusion matrix of the GB model trained on 50% overlapping 1.2s windows

and “Foot jiggling”. These two gestures are indeed quite similar, another source of misclassification is the “No fidgeting” class. A possible source of confusion here is that the “No fidgeting” class isn’t perfectly representative of the reality as subjects were asked not to move during the quiet sitting periods. A subject might do small movements that should not be considered as fidgeting. It is also to be noted that the classes are not perfectly balanced, with more “No fidgeting” samples compared to the other classes for the multiclass models. It is the opposite for the binary models, 2/3 of the samples are labelled as fidgeting.

7 Conclusion

In this work, we showed initial results on detecting fidgeting activities while seated. The algorithm can get decent performances on the multiclass problem and better performances when compared to the baseline algorithm for the binary classification problem. While more work is still necessary, this does provide initial promising results. The main limitation of this work is the collected dataset which is too small and leads to models overfitting. Therefore, the next step is to collect more data by having subjects in real-life situations instead of following a predefined protocol. Other possible outlook includes performing dimensionality reduction algorithm to select only the most relevant features and extracting features from other sensors such as the gyroscope or the force sensing insole.

References

1. Lugade, V., Fortune, E., Morrow, M., Kaufman, K.: Validity of using tri-axial accelerometers to measure human movement—Part I: posture and movement detection. *Med. Eng. Phys.* **36** (2), 169–176 (2014)
2. Zhang, T., Tang, W., Sazonov, E.S.: Classification of posture and activities by using decision trees. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4353–4356. IEEE, August 2012
3. Ayumi, V.: Pose-based human action recognition with Extreme Gradient Boosting. In: 2016 IEEE Student Conference on Research and Development (SCOREd), pp. 1–5. IEEE, December 2016
4. Zhang, T., Klein, D.A., Walsh, T., Lu, J., Sazonov, E.S.: Android TWEETY—a wireless activity monitoring and biofeedback system designed for people with Anorexia Nervosa. In: 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1–5. IEEE, June 2014
5. el Achkar, C.M., Lenoble-Hoskovec, C., Paraschiv-Ionescu, A., Major, K., Büla, C., Aminian, K.: Instrumented shoes for activity classification in the elderly. *Gait Posture* **44**, 12–17 (2016)
6. Munguia Tapia, E.: Using machine learning for real-time activity recognition and estimation of energy expenditure. Doctoral dissertation, Massachusetts Institute of Technology (2008)
7. Moufawad El Achkar, C., Lenoble-Hoskovec, C., Paraschiv-Ionescu, A., Major, K., Büla, C., Aminian, K.: Physical behavior in older persons during daily life: insights from instrumented shoes. *Sensors* **16**(8), 1225 (2016)
8. Pincus, S.M.: Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci.* **88**(6), 2297–2301 (1991)



Matching Confidence Masks with Experts Annotations for Estimates of Chromosomal Copy Number Alterations

Jorge Muñoz-Minjares¹ , Yuriy S. Shmaliy¹, Tatiana Popova²,
and R. J. Perez-Chimal³

¹ Department of Electronics Engineering, Universidad de Guanajuato,
36885 Salamanca, GTO, Mexico

{ju.munozminjares,shmaliy}@ugto.mx

² Centre de Recherche, Institut Curie,

Department of UNITE 830 INSERM Unité de Génétique et Biologie des Cancers,
75248 Paris, France

Tatiana.Popova@curie.fr

³ Department of Mechatronics, Universidad Tecnológica de Salamanca,
36766 Salamanca, GTO, Mexico

rperez@utsalamanca.edu.mx

Abstract. Structural aberrations (SAs), gains or losses in large segments of genomes, are associated with several genetic disorders. The SAs are commonly called the copy number alterations (CNAs) and their identification/classification is required to identify diseases. Many methods have been proposed to estimate the breakpoints and segmental constants in the CNAs with highest precision using the most powerful technologies of hybridization. However, locations and lengths of CNAs estimated using well-elaborated methods are often contradictory due to extensive variability of measurements and performance of the algorithms. Still much less attention is given to the estimation accuracy and it is difficult to select the best estimator. In this work, we propose to modify the confidence masks replacing the skew Laplace distribution with the asymmetric exponential power distribution (AEP) to approximate the *jitter* distribution in CNAs. Next, the estimates obtained using different algorithms are matched with the annotations made by experts employing the improved masks. Finally, we specify the match confidence probability of each CNAs detector algorithm respect the experts estimates.

Keywords: Copy number alterations · Modified confidence masks
Asymmetric exponential power distribution · Experts' annotations

1 Introduction

Somatic aberrations are abnormal changes in DNA called copy number alterations (CNAs) associated with diverse genetic disorders, mainly with cancer

disease [1,2]. Nowadays, a vast amount of technologies have been developed to measure the genome chromosomal structure: Array Comparative Genomic Hybridization (aCGH) [3], High Resolution CGH (HRCGH) [4], and Whole Genome Sequencing (WGS) [5] are among the most known. Nevertheless, the CNAs data obtained using the microarrays and other technologies are still affected by several factors: (1) nature of biological material (tumor is contaminated by normal tissue, relative values and unknown baseline for copy number estimation), (2) technological biases (quality of material and hybridization/sequencing) and (3) intensive random noise [6–8]. Consequently, intensive variability in measurements make an estimator, optimal or robust, unable to produce a reliable estimate [9]. Although the identification of CNAs associated with cancer is very challenging, still no one estimator can guarantee an existence of the detected changes.

2 Methods

To guarantee that an estimator has detected chromosomal changes accurately, the interpretation made by medical experts has been considered in [10] as a *gold standard*, although it has been recognized in [11] that such opinions are time consuming and not necessarily very accurate. Therefore, as has been shown in [8,12,14], testing estimates by confidence masks becomes challenging for medical applications. The masks allow improving the estimates for the required confidence probability by removing some CNAs, which do not match annotations made by experts.

2.1 Breakpoints Annotations

The annotations are counts of breakpoints in genomic regions made by visual inspection of the noisy signal. Observing each region, expert biologists determine whether or not it contains a breakpoint based on their expertise. Let us notice that visual annotations have been used successfully for object recognition in photos and cell phenotype recognition in microscopy [15,16]. An example of annotations for the breakpoints is shown in Fig. 1 as related to the Profile 44–Chromosome 1 taken from aCGH microarray experiments on neuroblastoma tumors. Here, data are separated on several regions, which were annotated by the experts as having “>0 breakpoints,” “1 breakpoint,” and “0 breakpoints.” Although the confidence probability was not specified, it can easily be deduced that the probability is high in each annotation.

2.2 Modified Confidence Masks

To guarantee an existence of CNAs with a required probability, an efficient algorithm computing the upper and lower boundary confidence masks has been proposed in [12]. The method is based on using the skew Laplace (SkL) law representing the jitter distribution in the CNA breakpoints. Later, this algorithm was

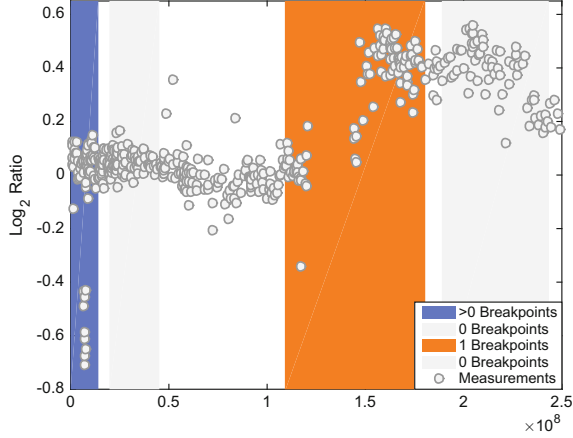


Fig. 1. Annotations made by medical experts to Profile 44–Chromosome 1 of sample of neuroblastoma. Regions with different annotations are separated. Experts suggested that there exist “>0 breakpoints,” “1 breakpoint,” and “0 breakpoints.” Data were obtained using the aCGH microarrays and plotted as Log_2 Ratio versus genomic position.

essentially improved referring to the fact that the Laplace distribution becomes highly inaccurate when the segmental signal-to-noise ratio (SNR) ranges below unity [17]. With respect to the l th breakpoint, the segmental SNRs are calculated as

$$\gamma_l^- = \frac{\Delta_l^2}{\sigma_l^2}, \gamma_l^+ = \frac{\Delta_{l+1}^2}{\sigma_{l+1}^2} \quad (1)$$

where $\Delta_l = a_{(l+1)} - a_l$, is the segmental difference and σ_l^2 and σ_{l+1}^2 are the segmental variances, which correspond to the measurements to the left l -segment and right $(l + 1)$ -segment.

The conclusion that has been made in [18] suggests that SkL is reasonably accurate when $\{\gamma_l^-, \gamma_l^+\} > 1$. The approximation error increases with $\{\gamma_l^-, \gamma_l^+\} < 1$ and can become unacceptable if $\{\gamma_l^-, \gamma_l^+\} \ll 1$. Therefore, It is needed to improve the jitter distribution based on the Asymmetric Exponential Power Distribution (AEPD).

Asymmetric Exponential Power Distribution. Following the procedure to measure the jitter pdf from [18], we arrive at the conclusion that it could be approximated with a sub-Laplacian distribution such as the AEP distribution [19], which is a generalization of the Gaussian and Laplace laws.

A random variable y_l associated with the l th breakpoint is said to have the AEPD, if for the shape parameter $\alpha_l > 0$, scale factor $\sigma_l > 0$, location $\theta_l = 0$, and skew factor $\kappa_l > 0$, a variable y_l is distributed with

$$p(k|\bar{p}_l, \bar{q}_l) = \frac{\alpha_l}{\sigma_l \Gamma\left(\frac{1}{\alpha_l}\right)} \frac{\kappa_l}{1 + \kappa_l^2} \begin{cases} \bar{p}_l^{k\alpha_l}, & k \geq 0, \\ \bar{q}_l^{|k|\alpha_l}, & k \leq 0, \end{cases} \quad (2)$$

where $\bar{p}_l = e^{-\frac{\kappa_l \alpha_l}{\sigma_l}}$, $\bar{q}_l = e^{-\frac{1}{\kappa_l \alpha_l \sigma_l}}$, and $\Gamma(x)$ is the Gamma function.

In a special case of $\kappa_l \neq 1$, letting $\alpha_l = 1$ transforms (2) to the discrete skew Laplace distribution [19], which alternative form is

$$p(k|\bar{p}_l, \bar{q}_l)_{\alpha_l=1} = \frac{1}{\sigma_l} \frac{\kappa_l}{1 + \kappa_l^2} \begin{cases} \bar{p}_l^k, & k \geq 0, \\ \bar{q}_l^{|k|}, & k \leq 0. \end{cases} \quad (3)$$

An example of applications of (2) as an approximation is given in [20], where factors α_l , κ_l and σ_l were estimated in the maximum likelihood sense for the simulated growth distribution. We found the constants α_l , κ_l and σ_l by fitting the histograms with a highest accuracy by minimizing the Kolmogorov–Smirnov distance [21] defined as

$$d_{KS} = \max|F_0(x) - S_N(x)|, \quad (4)$$

where $F_0(x)$ is the population cumulative distribution of (2) and $S_N(x)$ is the observed cumulative step function.

We calculate d_{KS} by (5) and select the minimum one to set the most appropriate values of α_l , κ_l , and σ_l for various symmetric SNRs $\gamma_l^- = \gamma_l^+$. The results can be found in Table 1. We also approximate α_l and σ_l in the mean square error (MSE) sense as

$$\alpha_l(\gamma_l) = 1 - \frac{a_1}{\gamma_l^{b_1}}, \quad (5)$$

$$\sigma_l(\gamma_l) = a_2 \gamma_l^{b_2}, \quad (6)$$

where $a_1 = 0.389$, $b_1 = 0.1394$, $a_2 = 1.142$ and $b_2 = -0.6289$. Note that, for $\gamma_l^- \neq \gamma_l^+$, the shape and scale factors are provided by $\alpha_l(\gamma_l^\pm) = \frac{\alpha_l(\gamma_l^+) + \alpha_l(\gamma_l^-)}{2}$ and $\sigma_l(\gamma_l^\pm) = \frac{\sigma_l(\gamma_l^+) + \sigma_l(\gamma_l^-)}{2}$. To find κ_l , we modified the equation proposed in [19]

$$\kappa_l = \left[\frac{\bar{X}_{\alpha_l}^-}{\bar{X}_{\alpha_l}^+} \right]^{\frac{1}{2(\alpha_l+1)}} \quad (7)$$

by substituting $\bar{X}_{\alpha_l}^- = \gamma_l^+$ and $\bar{X}_{\alpha_l}^+ = \gamma_l^-$ for the asymmetric case and $\kappa_l = 1$ otherwise. Several simulated measurements for equal SNRs and its approximations are showed in Fig. 2 taken from [22].

Table 1. Approximation MSEs produced by the SkL distribution and AEPD (2) with α_l , κ_l , and σ_l for different $\gamma_l^- = \gamma_l^+$.

$\gamma_l^- = \gamma_l^+$	α_l	κ_l	σ_l	SkL	AEPD
0.1	0.46	1	4.75	7.28×10^{-05}	1.11×10^{-07}
0.2	0.52	1	3.33	7.62×10^{-05}	7.09×10^{-08}
0.3	0.54	1	2.49	7.53×10^{-05}	1.16×10^{-07}
0.4	0.56	1	2.08	7.16×10^{-05}	1.03×10^{-07}
0.5	0.57	1	1.77	6.81×10^{-05}	1.24×10^{-07}
0.6	0.58	1	1.56	6.45×10^{-05}	1.04×10^{-07}
0.7	0.59	1	1.43	5.9×10^{-05}	8.33×10^{-08}
0.8	0.6	1	1.3	5.68×10^{-05}	9.99×10^{-08}
0.9	0.602	1	1.18	5.24×10^{-05}	1.13×10^{-07}
1.0	0.61	1	1.11	5.03×10^{-05}	1.21×10^{-07}
1.1	0.614	1	1.04	4.71×10^{-05}	1.13×10^{-07}
1.2	0.622	1	0.99	4.41×10^{-05}	7.07×10^{-08}
1.37	0.628	1	0.9	3.90×10^{-05}	6.80×10^{-08}
1.5	0.635	1	0.85	3.58×10^{-05}	6.32×10^{-08}
2.0	0.647	1	0.7	2.62×10^{-05}	7.45×10^{-08}

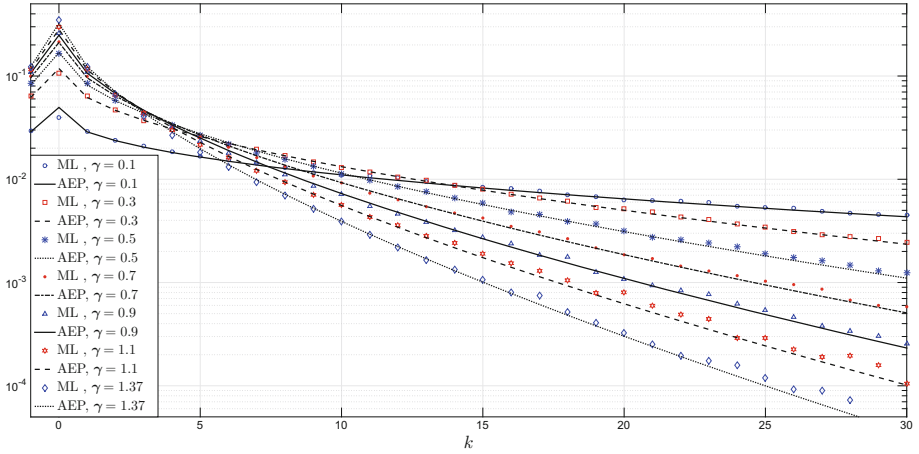


Fig. 2. Experimentally obtained one-sided jitter probability densities (markers) of the breakpoint location for equal SNRs in the range of $M = 800$ points with a true breakpoint at $n = 400$. The experimental density functions were found using the ML estimator. The histogram was plotted over 10^5 runs repeated 5 times and averaged.

Confidence Masks with AEP Distribution. The confidence masks are designed using statistical properties of the segmental noise and the breakpoint jitter to sketch the upper and low boundaries for the possible estimates.

For the AEPD-based approximation (2), the confidence masks can easily be modified using the equations given in [12] for the SkL distribution, in which case J_l^L and J_l^R can be defined if to specify $k_l^R(\vartheta)$ and $k_l^L(\vartheta)$ as [14]

$$k_l^R = \left\lfloor \frac{\nu_l \ln(1-p_l)(1-q_l)}{\kappa_l \xi(1-p_l q_l)} \right\rfloor, \quad (8)$$

$$k_l^L = \left\lfloor \nu_l \kappa_l \ln \frac{(1-p_l)(1-q_l)}{\xi(1-p_l q_l)} \right\rfloor, \quad (9)$$

where $\lfloor x \rfloor$ means a maximum integer lower than or equal to x . Note that functions (8) and (9) were obtained in [14] by equating (2) to $\xi(N_l) = \text{erfc}(\vartheta/\sqrt{2})$ and solving for k_l .

For the AEPD-based approximation (2), the UB and LB can be formed by replacing p_l and q_l with, respectively,

$$\bar{p}_l = e^{-\frac{\kappa_l^{\alpha_l}}{\sigma_l^{\alpha_l}}}, \quad (10)$$

$$\bar{q}_l = e^{-\frac{1}{\kappa_l^{\alpha_l} \sigma_l^{\alpha_l}}}. \quad (11)$$

That allows specifying the right-hand jitter \bar{k}_l^R and the left-hand jitter \bar{k}_l^L with, respectively,

$$\bar{k}_l^R = \left\lfloor \frac{\sigma_l \ln(1-\bar{p}_l)(1-\bar{q})}{\kappa_l \xi(1-\bar{p}_l \bar{q}_l)} \right\rfloor, \quad (12)$$

$$\bar{k}_l^L = \left\lfloor \sigma_l \kappa_l \ln \frac{(1-\bar{p}_l)(1-\bar{q}_l)}{\xi(1-\bar{p}_l \bar{q}_l)} \right\rfloor. \quad (13)$$

Provided (12) and (13), the jitter left boundary \bar{J}_l^L and right boundary \bar{J}_l^R can finally be defined as, respectively,

$$\bar{J}_l^L \cong \hat{n}_l - \bar{k}_l^R, \quad (14)$$

$$\bar{J}_l^R \cong \hat{n}_l + \bar{k}_l^L, \quad (15)$$

and substituted in the algorithm earlier designed in [12] for the SkL-based confidence masks. The masks can be formalized for the required probability, which will be associated below with the expert annotations.

3 Application

To specify the probability for the confidence masks, data used below were taken from the database of 575 annotated neuroblastoma copy number profiles, which are available from the public benchmark for testing new algorithms. We apply the algorithms to detect the CNA breakpoints in Chromosome 9 of Profile 1.

The BINSEG, PELT, SEGNEIGH and AMOC algorithms use a penalty value of $0.05 * \log(n)$, where n is the length of probes in the Chromosome 9, while the CBS algorithm detect the breakpoints automatically. The confidence masks applied are based on the asymmetric exponential power distribution with the probability ranging from 0.5 to 1. The detected CNA breakpoints, the SNR levels and the expert’s probabilities listed in (Table 2 and Fig. 3) illustrate the efficiency of the proposed masks. As can be seen, all estimators detect the first breakpoint. However, the second breakpoint is detected only by BINSEG, PELT, SEGNEIGH and CBS algorithms, although the breakpoint detected by CBS has a different location. The experts have noticed only the existence of the first breakpoint shown in (Fig. 3a). The confidence masks, displayed in (Fig. 3b), suggest that the second breakpoint does not exist with the probability of $P = 1 - 4.11 \times 10^{-11}$. For the CBS estimate shown in (Fig. 3c), the second breakpoint unlikely exists with the probability less than $P = 1 - 9.66 \times 10^{-04}$. These probabilities should be considered as confidence for the relevant estimators to match the experts’ notations.

Table 2. The breakpoints noticed by experts and the relevant probabilities for the confidence masks.

Method	Breakpoint genomic position	Levels of SNR	Matched probability to expert annotations
BINSEG, PELT, SEGNEIGH	27440311	$\gamma^- = 10.5, \gamma^+ = 17.2$	$P = 1 - 4.11 \times 10^{-11}$
	125714134	$\gamma^- = 2.1, \gamma^+ = 1.5$	
CBS	28766431	$\gamma^- = 11.6, \gamma^+ = 14.7$	$P = 1 - 9.66 \times 10^{-04}$
	136658606	$\gamma^- = 1.7, \gamma^+ = 1.4$	
AMOC	27440311	$\gamma^- = 11.6, \gamma^+ = 11.9$	–

4 Results

Estimates of the CNAs produced by different algorithms are often inconsistent due to various factors affecting probing. The confidence masks algorithm is appropriate to analyze the CNA measurements because the disturbances can be modeled as Gaussian noise. Accuracy of the CNAs estimates depends mostly on two factors: (1) segmental SNRs and (2) probe data length. For $SNR \gg 1$, the breakpoints can be identified in a visual way; therefore, the masks upper and lower boundaries computed for the skew Laplace distribution are highly accurate.

Otherwise, when $SNR \ll 1$, the CNA identification faces difficulties, both by experts and the masks, and the estimators may produce low confidence results. In this work, we have shown that the expert biologists are unable to identify the breakpoints with the confidence probability lesser than 99.99%. However, if expert’s probability is taken as the trustworthy standard, then the masks will

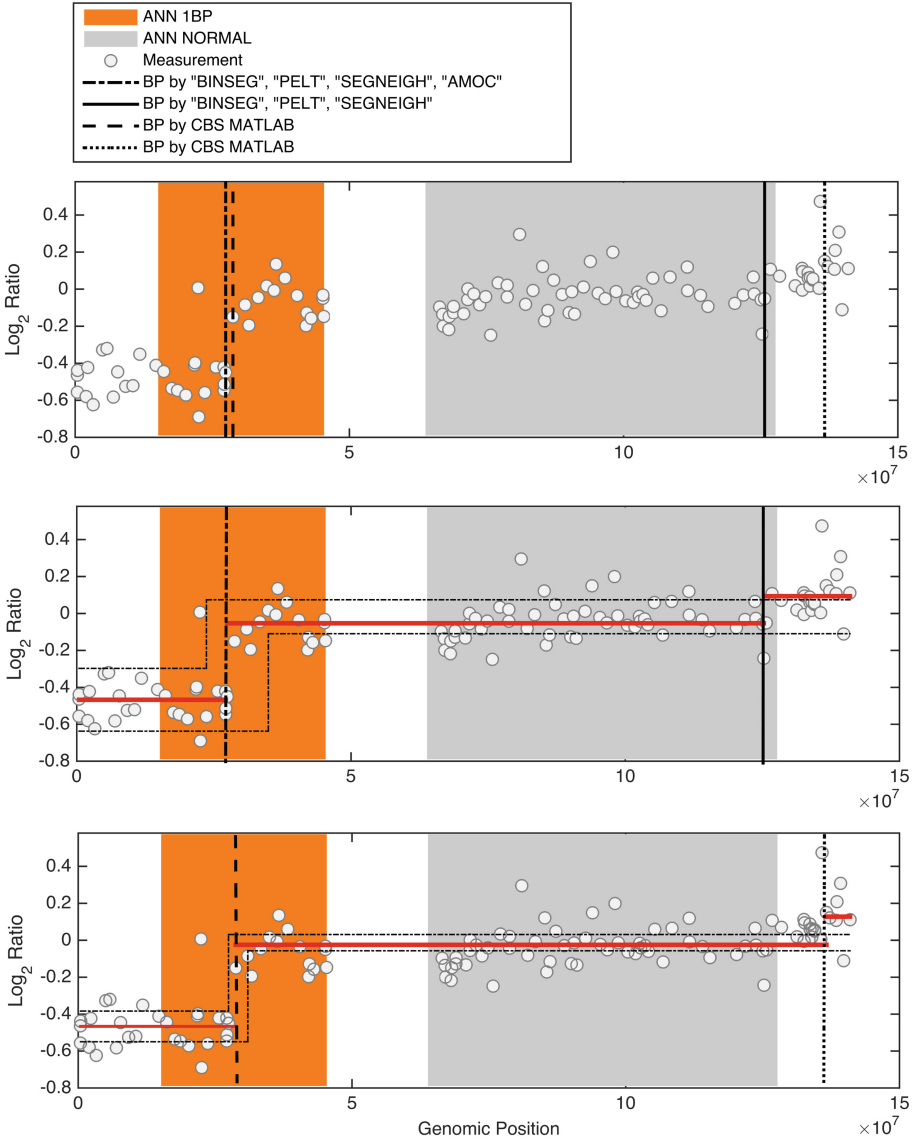


Fig. 3. Analysis of annotation regions made by experts and breakpoints found by several algorithms using the confidence masks. (a) Breakpoints found by several algorithms: BINSEG, PELT, SEGNEIGH, AMOC and CBS and two annotation regions one breakpoint 1BP and normal. (b) By applying the confidence masks to breakpoints found by BINSEG, PELT and SEGNEIGH, the second breakpoint is removed at a probability of $P = 1 - 4.11 \times 10^{-11}$. (c). The second breakpoint found by CBS algorithm unlikely exist to a lesser probability of $P = 1 - 9.66 \times 10^{-04}$ [13].

be able to make conclusions about the estimates matching the experts' opinion. Note that much subtler chromosomal changes can be efficiently examined using the masks with a smaller confidence probability.

5 Discussion

The confidence masks are intended to revise the CNAs estimates with the confidence probability matching annotations of medical experts. The propose method accounts for two factors: errors in the confidence masks and the database size. Firstly, the confidence masks still require more accurate approximations for the jitter distribution in the CNAs breakpoints. Because AEP distribution has been computed heuristically to fit the jitter distribution, an accurate mathematical model is required to decrease errors. Secondly, an acceptable confidence probability can be justified using an extensive number of annotations made by different qualified experts.

6 Conclusion

The AEPD has appeared to be more accurate than the discrete skew Laplace distribution to approximate the jitter distribution in the breakpoints of the CNAs. This holds particularly true for low and extra low Signal-to-Noise-Ratio values, $\text{SNR} < 1$ or $\text{SNR} \ll 1$, often observed in probes of small CNAs using CGH microarrays. The improved confidence probabilistic masks with the AEP approximation give a more correct picture for possible locations of CNAs on a probabilistic field.

Respect to the matched probability to experts annotations, we can deduce for this example that the AMOC method is the best to be used by these biologists, followed by CBS and finally by BINSEG, PELT and SEGNEIGH algorithms.

Solving problems mentioned in Discussion may open new horizons in further improvements of the CNAs estimates produced by different methods based on confidence probability of experts. The probabilistic confidence masks may play a crucial role in detecting actual chromosomal changes.

References

1. Graham, N.A., Minasyan, A., Lomova, A., Cass, A., Balanis, N.G., et al.: Recurrent patterns of DNA copy number alterations in tumors reflect metabolic selection pressures. *Mol. Syst. Biol.* **13**, 914 (2017)
2. Weinberg, R.A.: *The Biology of Cancer*. Gallard Science Taylor & Francis Group, LLC., London (2007)
3. Forozan, F., Karhu, R., Kononen, J., Kallioniemi, A., Kallioniemi, O.P., et al.: Genome screening by comparative genomic hybridization. *Trends Genet.* **13**, 405–409 (1997)
4. Speicher, M.R., Carter, N.P.: The new cytogenetics: blurring the boundaries with molecular biology. *Nat. Rev. Genet.* **6**, 782–792 (2005)

5. Ng, P.C., Kirkness, E.F.: Whole genome sequencing. *Methods Mol. Biol.* **628**, 215–226 (2010)
6. Zare, F., Dow, M., Monteleone, N., Hosny, A., Nabavi, S., et al.: An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. *BMC Bioinform.* **18**, 286 (2017)
7. Popova, T., Boeva, V., Manie, E., Rozenholc, Y., Barillot, E., et al.: Analysis of somatic alterations in cancer genome: from SNP arrays to next generation sequencing. In: *Sequence and Genome Analysis I Humans, Animals and Plants*. Ltd IP (edn.) iConcept Press Ltd. (2013)
8. Munoz, J.U., Shmaliy, Y.S.: Estimates of the breakpoints in genome copy number alteration profiles with masks. *Biomed. Sig. Process Contr.* **10**, 238–248 (2017)
9. Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., et al.: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.* **20**, 207–211 (1998)
10. Hocking, T.D., Schleiermacher, G., Janoueix-Lerosey, I., Boeva, V., Cappel, J., et al.: Learning smoothing models of copy number profiles using breakpoint annotations. *BMC Bioinform.* **14**, 164 (2013)
11. Tibshirani, R., Wang, P.: Spatial smoothing and hotspot detection for CGH data using the fused lasso. *Biostatistics* **9**, 18–29 (2007)
12. Munoz, J.U., Cabal, J., Shmaliy, Y.S.: Confidence masks for genome DNA copy number variations in applications to HR-CGH array measurements. *Biomed. Sig. Process Control* **13**, 337–344 (2014)
13. Minjares, J.M., Shmaliy, Y.S.: Matching confidence masks with experts annotations for estimates of chromosomal copy number alterations. *J. Genet. Disord.* **1**(1), 9 (2017)
14. Munoz, J.U., Shmaliy, Y.S., Cabal, J.: Confidence limits for genome DNA copy number variations in HR-CGH array measurements. *Biomed. Sig. Process. Control* **10**, 166–173 (2014)
15. Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T.: LabelMe: a database and web-based tool for image annotation. *Int. J. Comput. Vis.* **77**(1–3), 157–173 (2008)
16. Jones, T.R., Carpenter, A.E., Lamprecht, M.R., Moffat, J., Silver, S.J., Grenier, J.K., Castoreno, A.B., Eggert, U.S., Root, D.E., Golland, P., Sabatini, D.M.: Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc. Natl. Acad. Sci.* **106**(6), 1826–1831 (2009)
17. Munoz, J.U., Shmaliy, Y.S.: Improving estimates of genome CNVs with confidence masks using SNP array data. *Biomed. Sig. Process. Control* **31**, 238–248 (2017)
18. Munoz-Minjares, J.U., Cabal, J., Shmaliy, Y.S.: Effect of noise on estimate bounds for genome DNA structural changes. *WSEAS Trans. Biol. Biomed.* **11**, 52–61 (2014)
19. Ayebo, A., Kozubowski, T.J.: An asymmetric generalization of Gaussian and Laplace laws. *J. Prob. Statist. Sci.* **1**(2), 187–210 (2003)
20. Buldyrev, S.V., Growiec, J., Riccaboni, M., Stanley, H.E.: The growth of business firms: facts and theory. *J. Eur. Econ. Assoc.* **5**(2–3), 574–584 (2007)
21. Massey, F.J.: The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **46**(253), 68–78 (1951)
22. Munoz-Minjares, J., Shmaliy, Y.S., Olivera-Reyna, R., Olivera-Reyna, R., Perez-Chimal, R.J.: Jitter representation in SCNA breakpoints using asymmetric exponential power distribution. In: *14th International Conference on Electrical Engineering, Computing Science and Automatic Control* (2017)



Using Orientation Sensors to Control a FES System for Upper-Limb Motor Rehabilitation

Andrés F. Ruíz-Olaya¹, Alberto López-Delis^{2,3},
and Adson Ferreira da Rocha³

¹ Bioengineering Research Group, Antonio Nariño University,
Bogotá, Colombia

andresru@uan.edu.co

² Medical Biophysics Center, University of Oriente, Santiago de Cuba, Cuba
lopez.delis69@gmail.com

³ Bioengineering Group, University of Brasilia, Brasilia, Brazil

Abstract. Contralaterally controlled functional electrical stimulation (CCFES) is a recent therapy aimed at improving the recovery of impaired limbs after stroke. For hemiplegic patients, CCFES uses a control signal from the non-impaired side of the body to regulate the intensity of electrical stimulation delivered to the affected muscles of the homologous limb on the opposite side of the body. CCFES permits an artificial muscular contraction synchronized with the patient's intentionality to carry out functional tasks, which is a way to enhance neuroplasticity and to promote motor learning. This work presents an upper extremity motor rehabilitation system based on CCFES, using orientation sensors for control. Thus, the stimulation intensity (current amplitude) delivered to the paretic extremity is proportional to the degree of joint amplitude of the unaffected extremity. The implemented controller uses a control strategy that allows the delivered electrical stimulation intensity, to be comparable to the magnitude of movement. It was carried out a set of experiments to validate the overall system, for executing five bilateral mirror movements that include human wrist and elbow joints. Obtained results showed that movements voluntary signals acquired from right upper-limb were replicated successfully on left upper-limb using the FES system.

Keywords: Motor rehabilitation · Joint angle estimation · Orientation sensors
Contralaterally controlled functional electrical stimulation

1 Introduction

Motor disorders of the upper limb are a kind of effects occurring after stroke, affecting the execution of movement, leading to lack of coordination and motor control. Stroke is a significant cause of motor disability in most of the countries around the world, affecting people to execute activities of daily living (ADL) [1]. Specific rehabilitation therapies may provide functional motor recovery. Rehabilitation of motor function has been linked to motor learning that occurs during repetitive, frequent and intensive training. This sensorimotor activity base on neural plasticity, which is the ability of health areas of the brain to reorganize and compensate for lost function in other brain

regions [2]. The human neuromuscular system exhibits use-dependent plasticity, which means that use modifies the properties of neurons and muscles, their connectivity, and thus their function. Repetitive motor activity in a real-world environment with a cognitive effort has been identified in several studies as favorable for motor recovery of post-stroke patients [2]. In literature, functional electrical stimulation (FES) has been evaluated to assist motor recovery processes in people after stroke [3, 4]. FES is a technology that permits to generate controlled specific muscle functions using electrical impulses [5, 6]. As exposed above, it is required to directly engage the patient with the functional task by actively initiating or controlling electrical stimulation under voluntary control. Contralaterally controlled functional electrical stimulation (CCFES) is a recent therapy aimed at improving the recovery of impaired limbs after a stroke. This technique uses a voluntary control signal from the non-impaired side of the body to regulate the intensity of electrical stimulation delivered to the impaired muscles of the opposite side of the body [7, 8]. For example, during therapy of hemiplegic people with hand CCFES, the stimulation intensity delivered to paretic hand is proportional to the degree of movement of the unaffected hand [7].

The CCFES therapy incorporates several rehabilitation principles that are important for motor relearning. First, CCFES puts the brain back in control of the affected limb by giving the patient direct proportional control of the stimulation intensity, [7]. With CCFES, the patient not only controls when the stimulation turns on (as is also the case with EMG-triggered or switch-triggered stimulus) but also manages the duration and intensity of stimulation and therefore the resultant movement [7]. Second, CCFES incorporates a mirroring strategy. It consists of synchronous mimicking of the behavior of the healthy limb, by having the contralateral unaffected limb control stimulation and consequent movement of the affected limb. Third, CCFES creates a high perception of restored motor control because the stimulated arm moves according to the patient's motor intention [7, 8].

CCFES requires detecting motion intention. This detection requires technical issues concerning, for example, discrete signal extraction (either from the brain or the muscles), noise deletion, etc. Motion intention detection can be realized through motion sensors, force/torque sensors, electromyography, or electroencephalography. This work obtains the control signal from the non-paretic side of the subject from orientation sensors. Thus, the stimulation intensity (e.g., current pulse width) delivered to the paretic extremity is proportional to the degree of joint amplitude of the unaffected extremity. Orientation sensors are a kind of wearable sensors that are used to gather movement data. These devices base on the integration of accelerometers, gyroscopes, and magnetometers. Unlike an inertial measurement unit (IMU), which only provides raw sensor readings, the orientation sensor includes an onboard microcontroller that combines sensor data using specific sensor fusion algorithms mainly based on Kalman filtering to generate orientation estimates [9, 10]. The combination of multiples orientation sensors placed on connected body segments and the additional information on the kinematic constraints enable joint angles to be measured in an adequate way [10]. This work describes the implementation of a CCFES platform for upper-limb motor rehabilitation that uses orientation sensors for detecting motion intention. It was carried out a set of experiments for executing five bilateral mirror movements that include human wrist and elbow joints.

This paper is organized as follows. Next section presents the developed system that uses control signals from orientation sensors. Following is described the experimental methods including the protocol, data processing, validation, and results. Finally, the last section presents conclusions and future work.

2 Orientation Sensor-Based CCFES System

2.1 Architecture of the System

Figure 1 shows the upper-limb motor rehabilitation platform to implement CCFES.

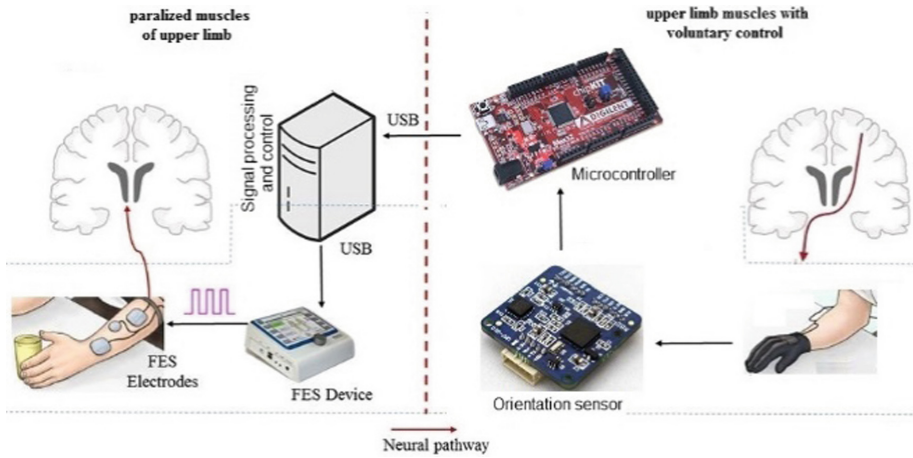


Fig. 1. Upper-limb motor rehabilitation system to implement orientation sensor-based CCFES.

The figure above represents an upper-limb closed-loop system based on the CCFES paradigm. The implemented system includes two human joints: wrist and elbow. The joint amplitude of the unaffected extremity limb was detected using orientation sensors and processed to control the electrical stimulation intensity of muscles in paralyzed limb, to generate mirror movements with the unaffected arm. The above offers the user the ability to control the stimulation involving a cognitive task.

2.2 Joint Amplitude Estimation

It was used a sensor that provides orientation estimates concerning a fixed reference frame which is formed by the gravity vector and the Earth's magnetic North vector, regarding Euler angles. Thus, the sensor estimates the rotation of the body part from the device is placed on, relative to a fixed point on the subject's body (see Fig. 2).

A joint rotation is defined as the orientation of a distal segment (L1) concerning a proximal section (L2). In the biomechanical analysis, body parts are simplified as rigid bodies. Taking into account that it is required tracking joint angle of the human wrist

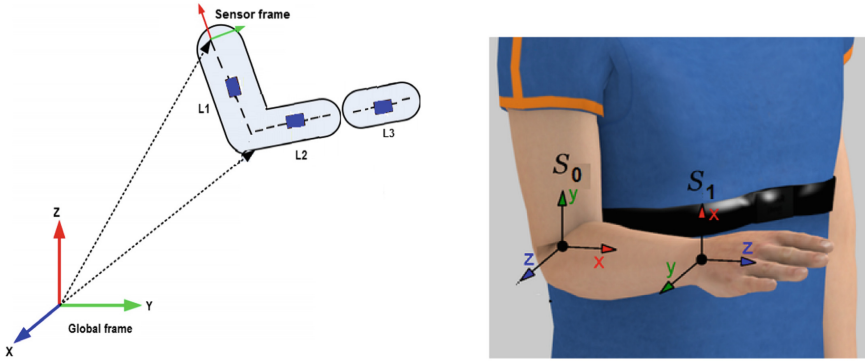


Fig. 2. Relation of the orientation sensor frame with global frame for the elbow and wrist joint. A joint rotation defined as the orientation of a distal segment with respect to a proximal segment.

and elbow joints, it was used a kinematic model consisting of a structure with three segments linked by revolute joints. The arm, attached to the shoulder and elbow joints; the forearm, attached to the elbow and wrist joints; and the section belonging to the hand; attached to the forearm by the wrist joint [11]. Orientation sensors were placed on the three segments.

For determining the position of each sensor in the reference (global) coordinate system, it is required to transform the inertial measurements from the sensor coordinate system to the reference (global) coordinate system. This paper uses a representation of the wrist and elbow joint angles based on rotation matrices.

Rotation matrices contain information about the relative position of two coordinate systems (regarding Euler angles) so that they can be used to transform any point in one coordinate system to another [11]. These matrices represent the sensor orientation information. Equation 1 presents the rotation matrix, where ψ is the rotation about the z-axis, θ is the rotation about the y-axis and φ is the rotation about the z-axis:

$$R = R_{\psi}^Z \cdot R_{\theta}^Y \cdot R_{\varphi}^Z = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix} \quad (1)$$

L1, L2, and L3 in Fig. 2 represent the subject anthropometric measurements of arm, forearm, and hand, respectively. Assuming that R_{GS} is the rotation matrix that rotates a vector in the sensor coordinate system (S) to the global reference system (G), then:

$$x_G = R_{GS} \cdot x_S \quad (2)$$

Taking into account the position of the orientation sensor relative to the segment it represents (upper arm or forearm), the reconstruction of the limb position is performed as follows:

$$arm = R_{GS}^A \cdot \begin{pmatrix} L1 \\ 0 \\ 0 \end{pmatrix} \quad forearm = R_{GS}^F \cdot \begin{pmatrix} L2 \\ 0 \\ 0 \end{pmatrix} \quad hand = R_{GS}^H \cdot \begin{pmatrix} 0 \\ -L3 \\ 0 \end{pmatrix} \quad (3)$$

R_{GS}^A , R_{GS}^F , and R_{GS}^H represent the 3×3 rotation matrices corresponding to Sensor1, Sensor2, and Sensor3, respectively. The vectors arm , $forearm$ and $hand$, correspond to the relative upper limb segments in the global frame coordinates. Wrist joint angle (Θ_w) and elbow joint angle (Θ_e) are obtained directly from both vectors.

2.3 FES Controller

In literature, there are several ways to implement a controller for a FES system, including open-loop, closed-loop and state machine controllers [12]. Open-loop controllers are relatively simple to implement, and consist of an electrical stimulator and a command source, which could be a switch controlled by the patient or therapist or a specific muscle activity.

In this work, it was implemented an open-loop controller that employs a proportional control strategy. This kind of controller allows the control of electrical stimulation intensity, to be comparable to the magnitude of voluntary signals. Specifically, as the voluntary information, it was used the degree of joint amplitude of the unaffected extremity (i.e., Θ_w and Θ_e). Thus, the level of users' muscle contraction response (im-paired limb) will be in proportion to the joint amplitude recorded from the human joint (non-impaired limb). With this controller, the user provides the inputs that are determined based on proprioceptive feedback, vision, and experience [13]. Figure 3 shows used relationship joint amplitude-electrical stimulation to implement the FES controller.

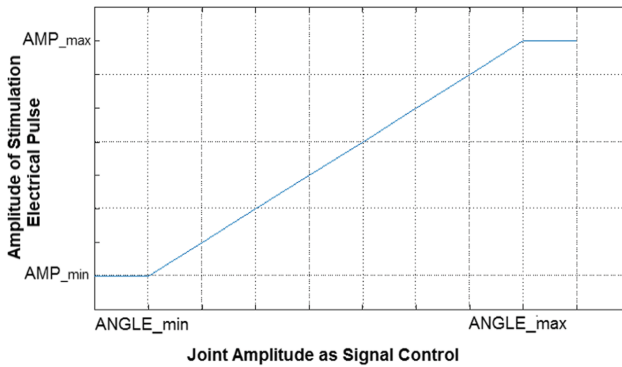


Fig. 3. Linear input-output relationship between the stimulation pulse and the joint amplitude as signal control.

AMP_min represents the amplitude of the stimulation pulse just before to generate a motor response. AMP_max represents the amplitude of the stimulation pulse to produce a maximum joint flexion or extension. ANGLE represents the value of signal control (Θ_w and Θ_e). When ANGLE signal was below the inferior threshold ANGLE_min, the controller delivers a stimulation pulse of amplitude AMP_min, on the contrary, when the ANGLE signal was above ANGLE_max the controller delivery a stimulation pulse of amplitude AMP_max. In others side, when the ANGLE signal was between two thresholds, the relationship was linear.

3 Experimental Methods

3.1 Experimental Setup and Protocol

A healthy volunteer subject participated under informed consent. It was carried out anthropometric measurements of upper limb segments. Three orientation sensors (model UM7-LT from CH Robotics) were coupled to the right arm, forearm, and hand. Each orientation sensor provides absolute angular position.

Signals from orientation sensors were acquired through 3 serial communication ports using a Chipkit Max32 with a sampling frequency of 50 Hz. Information on elbow joint angle and wrist joint angle was sent to control unit using serial communication (speed 115200).

In order to evaluate human joint amplitude estimation from orientation sensor data, it was used as reference a goniometer, which is an instrument used to measure angles within the field of physical therapy and occupational health. It was manually adjusted and set on arcs of 0, 30, 60, 90 and 120° for elbow flexo-extension, and it was set on arcs of 0, 30 and 60° for flexion and extension of the wrist.

3.1.1 Execution of Bilateral Mirror Movements

The FES equipment used to deliver electrical stimulation was the Rehastim model from Hasomed GmbH, which has eight channels to stimulate up to 8 muscles with attached electrodes. The Rehastim equipment was set in ‘science mode’ to enable control from a computer using Matlab-Simulink. It was connected four electrodes on the left upper limb: two arm muscles (biceps brachii long head and triceps brachii long head) which are involved in elbow flexion-extension movements; and two forearm muscles (flexor carpi radialis and extensor carpi radialis) which are involved in wrist flexion-extension movements, see Fig. 4.

A subject was seated in front of a computer that indicated movements to execute. In each trial, subject repeated each limb motion three times following five movements: flexion and extension of the wrist, flexion and extension of the elbow, and resting state. The order of these limb motions was randomized. The resting state was kept in both at the beginning and at the end of each test, as well as between movements. A total of three trials were complete in a session.



Fig. 4. Position of the electrodes, four electrodes were attached to the left upper limb: biceps brachii long head (top-left), triceps brachii long head (top-right), flexor carpi radialis (bottom-left) and extensor carpi radialis (bottom-right).

3.2 Data Acquisition and Processing from Platform Simulink

Figure 5 shows the Simulink model to the implementation electrical stimulation. There is a serial communication to receive information about joint angles from microcontroller (Fig. 5, top). This data was sent to a Simulink submodel named “Electrical Stimulation RehaStim”, which is detailed in Fig. 5, below.

A real-time synchronization block has been used to ensure a Simulink execution frequency of 20 Hz. It achieves this by synchronizing the Simulink FES control system with the real-time clock of the computer. A stimulator interface permits accessing the RehaStim stimulator.

The control parameters of the applied electrical pulses are the amplitude, pulse width, and frequency. The pulse width was set to 140 μ s, the frequency was set to 20 Hz, and the magnitude varies according to the control signal following relationship showed in Fig. 3. Selected thresholds from control signal were: ANGLE_min (5% of maximum joint flexion or extension) and ANGLE_max (95% of maximum joint flexion or extension). Table 1 presents the empirical approach used to obtain the parameters of AMP_min and AMP_max that depends on each stimulated muscle.

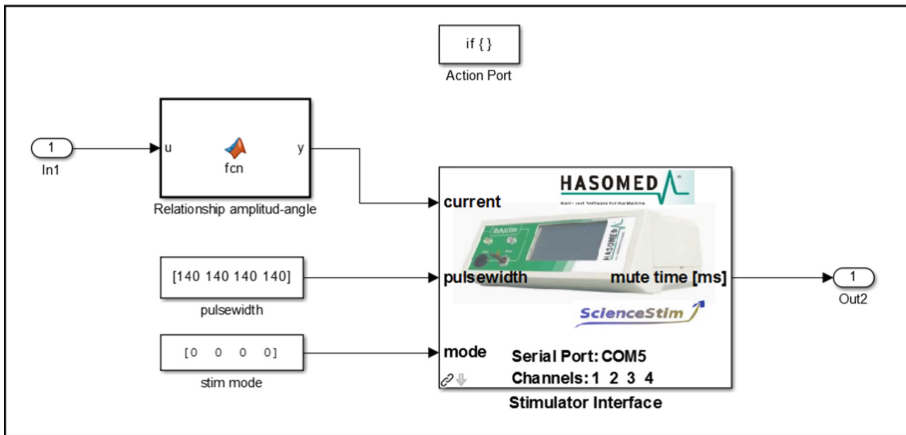
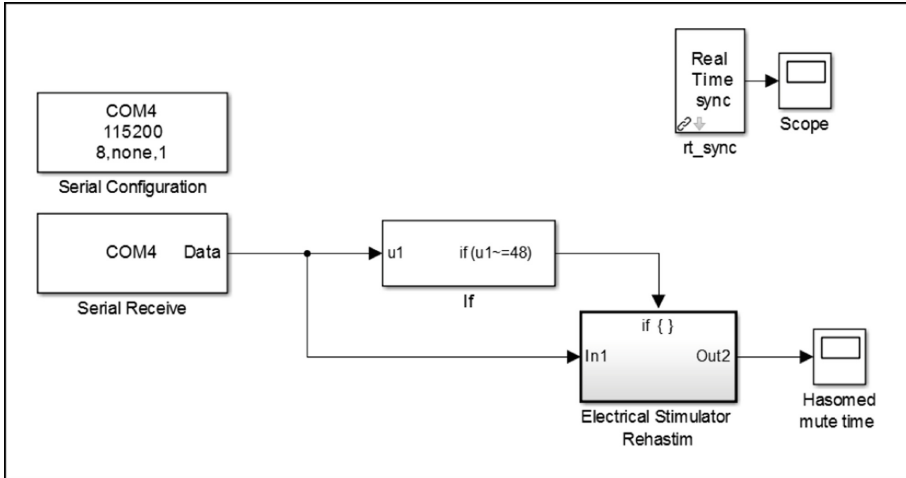


Fig. 5. Simulink model to control the rehastim electrical stimulator. Top: serial communication to receive information about joint angles from microcontroller. Bottom: FES controller

Table 1. Parameters of amplitude for electrical stimulation

Parameter	Upper limb muscles			
	Biceps bracci	Triceps bracci	Flexor carpi radialis	Extensor carpi radialis
AMP_min (mA)	11	10	8	8
AMP_max (mA)	23	21	16	15

3.3 Results

Figure 6 shows estimated joint angles (dotted lines) for 0, 30, 60, 90 and 120° for elbow flexo-extension. The root-mean-square error (RMSE), see Eq. 4, between the estimated values of orientation sensors and the goniometer values was calculated.

$$e(n) = \theta_{\text{goniometer}}(n) - \theta_{\text{sensor}}(n)$$

$$\text{RMSE} = \sqrt{\frac{1}{k} \sum_{n=1}^k e^2(n)} \quad (4)$$

The performance of both sensing devices (goniometer and orientation sensors) is quite similar, according to the total mean error calculated below to 5.3° for elbow joint angle and 5.8° for wrist joint angle. Those results showed that joint angle information from orientation sensors provide adequate accuracy to be used in the platform described in Fig. 1.

Regarding to the randomized movements executed by the subject (flexion and extension of the wrist, flexion and extension of the elbow, and rest), all the sequence of movements were replicated successfully on left upper-limb using the FES system.

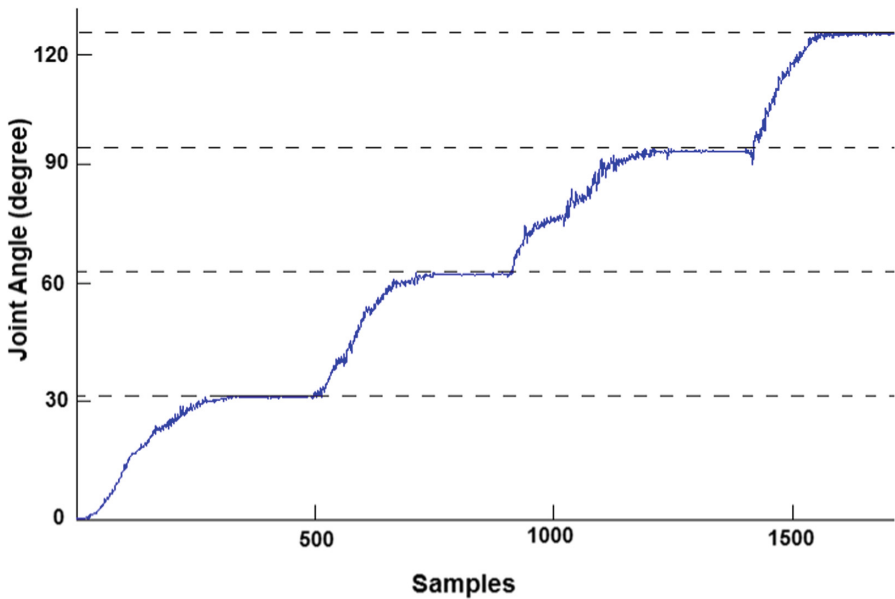


Fig. 6. Estimated joint angles for 0, 30, 60, 90 and 120° for elbow flexo-extension. The dotted lines represent each estimated angle from orientation sensors

4 Conclusions

The proposed FES system is a promising motor training aid for patients with upper limb impairment after stroke. The system permits an active FES-assistive therapy based on a novel motor functional rehabilitation idea, using control signal from orientation sensors adhered on a healthy limb to regulate the stimulus parameters of stimulation applied to the paralyzed limb, to achieve synchronous movement of bilateral limbs. The proposed system has several advantages for motor rehabilitation such as intention-driven movement, bilateral movement, and creating a high perception of restored motor control [14]. Obtained results showed that all five defined upper-limb voluntary movements executed from right upper-limb were replicated successfully on left upper-limb using the FES system.

Next work relates to implement a close loop controller. Thus, it is required: (1) to attach sensors to the impaired limb to measure the actual movement; (2) to obtain a more realistic input-output relationship between the stimulation pulse and the joint amplitude, than the linear relationship used.

Furthermore, as future work, incorporating therapy exercises into virtual games can motivate the patient to put in the effort and encouraging more exercises. Also, the use of virtual reality enables more realistic activity to be performed.

Acknowledgments. Authors want to thank to the Universidad Antonio Nariño in Colombia for supporting this work under project number 2016217 and the CNPq for productivity and post-doctoral fellowship.




References

1. World Report on Disability, World Health Organization (WHO) (2011)
2. Moller, A.R.: *Neural Plasticity and Disorders of the Nervous System*. Cambridge University Press, Cambridge (2006)
3. Hara, Y., Obayashi, S., Tsujiuchi, K., Muraoka, Y.: The effects of electromyography controlled functional electrical stimulation on upper extremity function and cortical perfusion in stroke patients. *Clin. Neurophysiol.* **124**, 2008–2015 (2013)
4. Sheffler, L., Chae, J.: Neuromuscular electrical stimulation in neurorehabilitation. *Muscle Nerve* **35**, 562–590 (2007)
5. Doucet, B.M., Lamb, A., Griffin, L.: Neuromuscular electrical stimulation for skeletal muscle function. *Yale J. Biol. Med.* **85**, 201–215 (2012)
6. Popovic, D.B., Sinkjær, T., Popovic, M.B.: Electrical stimulation as a means for achieving recovery of function in stroke patients. *NeuroRehabilitation* **25**, 45–58 (2009)
7. Knutson, J.S., Harley, M.Y., Hisel, T.Z., Makowski, N.S., Fu, M.J., Chae, J.: Contralaterally controlled functional electrical stimulation for stroke rehabilitation. In: *Proceedings of IEEE Engineering and Medicine and Biology Society*, pp. 314–317 (2012)
8. Knutson, J.S., Harley, M.Y., Hisel, T.Z., Makowski, N.S., Chae, J.: Contralaterally controlled functional electrical stimulation for recovery of elbow extension and hand opening after stroke: a pilot case series study. *Am. J. Phys. Med. Rehabil.* **93**(6), 528–539 (2014)

9. Sabatini, A.M.: Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing. *Sensors* **11**, 1489–1525 (2011)
10. Filippeschi, A., Schmitz, N., Miezal, M., Bleser, G., Ruffaldi, E., Stricker, D.: Survey of motion tracking methods based on inertial sensors: a focus on upper limb human motion. *Sensors* **17**, 1257 (2017)
11. Borbély, B.J., Szolgay, P.: Real-time inverse kinematics for the upper limb: a model-based algorithm using segment orientations. *Biomed. Eng. Online* **2017**(16), 21 (2017)
12. Lynch, C., Popovic, M.: Functional electrical stimulation: closed-loop control of induced muscle contractions. *IEEE Control Syst. Mag.* **28**, 40–49 (2008)
13. Ferrarin, M., Palazzo, F., Riener, R., Quintern, J.: Model-based control of FES-induced single joint movements. *IEEE Trans. Neural Syst. Rehabil. Eng.* **9**(3), 245–257 (2001)
14. Knutson, J.S., Gunzler, D.D., Wilson, R.D., Chae, J.: Contralaterally controlled functional electrical stimulation improves hand dexterity in chronic hemiparesis. *Stroke*. **47**(12), 2596–2602 (2016)



A Real-Time Research Platform for Intent Pattern Recognition: Implementation, Validation and Application

Andres F. Ruiz-Olaya¹(✉) , Gloria M. Díaz² , and Alberto López-Delis³ 

¹ Universidad Antonio Nariño, Bogotá, Colombia
andresru@uan.edu.co

² Instituto Tecnológico Metropolitano, Medellín, Colombia
gloriadiaz@itm.edu.co

³ Centro de Biofísica Médica, Universidad de Oriente, Santiago de Cuba, Cuba
alberto.lopez@cbiomed.cu

Abstract. Despite multiple advances with myoelectric control, currently there is still an important need to develop more effective methods for controlling prosthesis and exoskeletons in a natural way. This work describes the design and development of a research tool for the design, development and evaluation of algorithms of myoelectric control which base on intention detection from neuromuscular activation patterns. This platform provides integrated hardware and software tools for real-time acquisition, preprocessing, visualization, storage and analysis of biological signals. It is composed of a bio-instrumentation system controlled by a real-time software created in Simulink and executed on the xPC-target platform and, a Java based software application that allows to manage the acquisition and storage processes by a system operator. System evaluation was performed by the comparison with reference signals provided by a function generator and, as an example of the application of the developed acquisition platform, it was carried out a set of experiments to decode movements at the upper-limb level.

1 Introduction

Movement intention based on pattern recognition of electrophysiological signals are becoming increasingly important in many clinical and research applications, such as personalized health systems, rehabilitation applications, and active control of prosthesis devices. These technologies promise be an alternative for improving the quality of life of people with pathologies of the nervous system, amputations, and other physical disabilities [1, 2].

The use of pattern recognition approaches for identifying movement intention has been increasing in the last years. Actual approaches are based on the classification of signals that represent the anatomical and physiological properties

of neuromuscular activity during the muscle contraction, such as electromyography (EMG) [3]. EMG pattern recognition based control has emerged as a feasible alternative in rehabilitation robotic devices. The effectiveness of EMG-based classification and pattern recognition algorithms is based on effective implementation of three modules: pre-processing, feature extraction and pattern classification, which must to perform a real-time execution, so the computational cost should not be high [4].

Performance evaluation of those algorithms becoming a critical aspect in the pattern recognition research area, in which it is important to obtain the parameters characterization for highest classes separation, robustness, and complexity. To ensure the quality and validity of the performance reported in each particular study, researchers require collecting a set of real cases that are used for training and testing the proposed models and for evaluating different experimental setups. These experimental setups can include changing the information representation strategy (feature extraction) but also integrating or fusing other biological measures into the analysis [5].

This paper focuses on development and application of a Matlab-based experimental platform to provide a flexible prototyping environment to test myoelectric control algorithms for robotic prostheses and exoskeletons. The platform permits implementation of a variety of EMG-based algorithms for (1) Signal processing; (2) Feature extraction; (3) Pattern recognition; (4) Sensor data fusion, and, (5) Real-time myoelectric control. The flexible development approach allows the integration of new features, classifiers as well as the extension to other types of sensor data. The platform supports the real-time acquisition of case sets in a research environment. Data capturing can be performed by non-expert operators using a friendly graphic user interface that allows to recording sets of signals, specifically EMG and proprioceptive information, morphological measures and clinical data according to a study designed by the researcher. Signal acquisition parameters and preprocessing methods can be developed easily in Matlab/Simulink and to be used in the capturing process mean of the xPC-Target platform [6]. Finally, case sets can be used for evaluating statistical learning models developed offline in any computational tool. While existing toolboxes and software focus on purpose general time series processing, the experimental platform in this contribution has specifically been developed for intent pattern recognition.

A study of validation was carried out based on sensor data fusion. Thus, it was carried out a set of experiments to compare two EMG-based pattern recognition algorithms: one algorithm using EMG data and other algorithm using kinematics information and EMG data. This study aims to classify five movements at the upper-limb level. Results show that the use of combined EMG and proprioceptive signals improve the classification respect to use only EMG.

Paper is organized as follows: in the next section hardware and software system components are presented. Section 3 describes experimental methods to evaluate the acquisition platform. Section 4 presents obtained results in the validation stage, and the last section presents conclusions of this work.

2 System Components

A general diagram illustrating the main components of the developed system is presented in Fig. 1. As is schematized in this figure, the patient or subject study is connected to the bio-instrumentation hardware, which is composed of a set of superficial EMG and proprioceptive sensors. The real-time acquisition process is performed by an acquisition module based on the xPC-target component of Matlab, this module is controlled remotely by a java based graphical user interface, which also allows to manage the definition of experimental protocol and the information storage process.

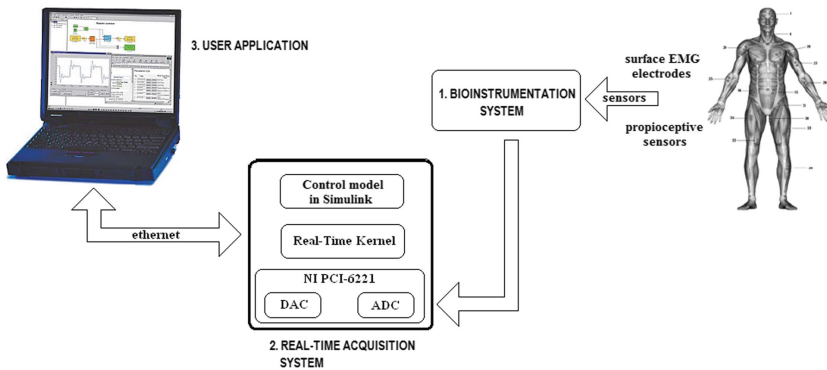


Fig. 1. Architecture of the proposed system composed of three modules: 1. Bioinstrumentation hardware, 2. xPC-target based acquisition system and 3. User application

2.1 Bio-instrumentation Hardware

The bioinstrumentation hardware allows acquiring bioelectric (EMG) and proprioceptive (velocity and acceleration of a body segment) signals. The software platform may be used jointly with a bioinstrumentation system to acquire data of several muscular groups. Since EMG signal is small ($50 \mu\text{V}$ – 5mV), acquired EMG signals may be affected by interference from other biological and environmental sources like movement artifacts, electric noise, muscle noise, etc. So, the raw EMG signals must be amplified by bioinstrumentation hardware to be acquired in the data acquisition board. Hardware implemented a band pass filter with a cut-off frequency between 20 and 500 Hz and provides amplification gains, which can be controlled from the software platform.

2.2 Real-Time Signal Acquisition System Based on xPC-Target

xPC-Target is an operating system developed by Mathworks that emulates a real-time system on a computer so that the whole system runs in real-time. It can be used for signal acquisition, rapid prototyping, testing, and deploying real-time systems [6]. A xPC-Target environment is composed of a host and a target

PCs; real-time C++ applications are developed into the host PC, using the real-time workshop Matlab/Simulink coder, and loaded to the target PC that runs the xPC-Target kernel. At runtime, the execution of target programs can be controlled by the host PC. Host-target communication is performed through the TCP/IP communication protocol in an Ethernet link [6].

The use of MATLAB/Simulink allows that the acquisition parameters such as gain amplification and signal filtering can be experimentally determined, simulated and validated in a rapid control prototyping (RCP) scheme. Due that the code is automatically generated by the Simulink coder, the researcher doesn't have to worry about the implementation of algorithms in the source code. In consequence time to testing different parameters and algorithms is reduced. Moreover, xPC-Target allows easy integration of future modules required for investigating patterns of neuromuscular activation such as on-line learning models and control strategies of prosthesis and orthosis based on biomechanical patterns.

2.3 User Application

The user application has been developed for allowing to non-expert users interacting with the data acquisition platform. Software architecture followed the well-known model-view-controller pattern design. The view layer, which includes all the panels and dialogs that allow the interaction with the user, was developed in Java-Swing API. The controller layer, includes the classes that manage the logic of application, it is composed by three main packages, the study/subject manager package, which controls the logic of data extracted from persistence modules; the signal acquisition package, which implements all process of signal acquisition with a low dependency of xPC-Target (less than 10% of the overall classes), aiming to change the acquisition model without require several software modifications; finally, the xPC target communication package, which is an interface package that implements all the calls to the xPC Target API to allows the real time signal acquisition. This package is the only with full dependency of xPC Target. The third layer implements overall persistence model mean both JDBC connection and the XML files.

Graphic User Interface (GUI). The GUI manages the display of signals and allows to capturing the biomedical data of subjects participating in each particular study. Functional requirements were provided by experimented researchers on biomechanical analysis and patterns neuromuscular activation, who highlighted the necessity of to acquire, store and retrieve both clinical data and biosignals, to provide tuning parameters of models developed in Simulink and to manage the signal capturing and recording processes. To integrate above characteristics, the next interfaces were developed:

Subject data management. This module provided operator to register, search and edit subject participating in any particular study. When a new subject is created in the database, a set of clinical data and morphometric measurements are also captured, which can be used for posterior analysis.

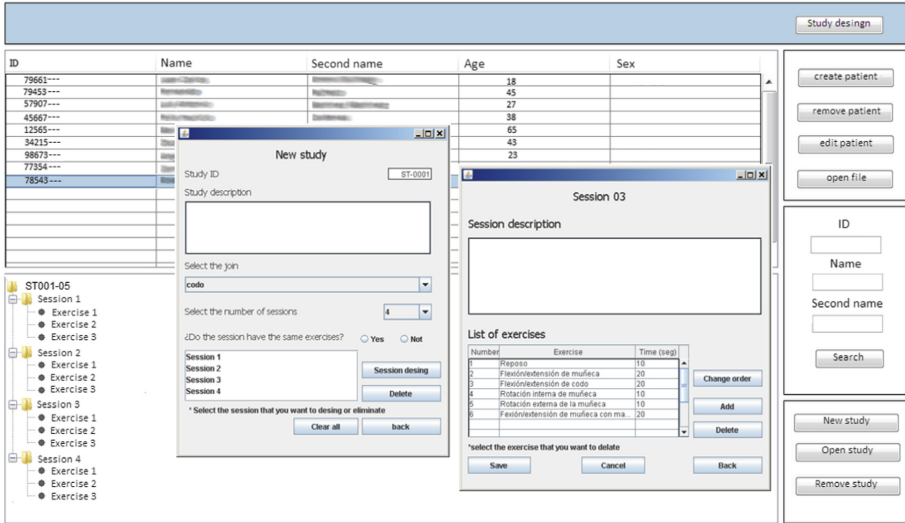


Fig. 2. Study and session design interfaces (front). Subject data management (back)

Study and session design. This panel provides researcher to define experimental protocols that allow to develop a specific study. A study is here defined as a set of one or more sessions each composed of a set of trials (movements or resting state). Once a study had been created, the number of sessions is defined, and each of them is designed defining a list of ordered trials, each endowed with a demonstration video, which will be displayed in the acquisition stage. Screens displaying the main panels of the Study and session design modules are shown in Fig. 2.

Signal acquisition and visualization. This module provided operator to capturing biological data defined by any particular designed study. Once the subject is registered, the operator starts the new study module, which requires a set of morphometric measurements that had been defined for each articulation. Then, data acquisition for any session can be carried out. As it is shown in the Fig. 3, the operator can select the signal channels and the signal Gain that will be used during the acquisition. Additionally, the signals that are actually captured are also displayed, to provide them to verify the acquisition system operation. Once the system has been configured, a training phase can be performed, so that the subject is familiar with the procedure, or the acquisition process can be started. In any case, the demonstration video that shows the sequential series of movement repetitions is presented to the subject (see Fig. 3, left). Stored signals can also be displayed for visual analysis from the visualization module, which provided user to manipulate aspects of signal displaying such as signal scale time, signal scale amplitude and signal scale gain.

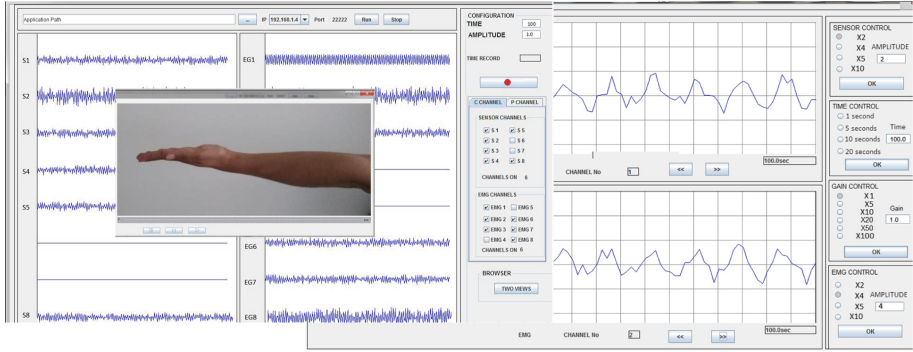


Fig. 3. Signal acquisition and visualization interfaces

xPC-Target Communication Module. Creating custom software to control a real-time application running on the target PC is feasible by the use of the xPC-Target API dynamic link library (DLL) or COM. To access the xPC-Target API from a Java environment, a Java wrapper based on Java Native Access (JNA) technology¹ was developed, which would allow xPC-Target API functions to be managed from any Java application in either Windows or Linux/Unix operative system. This module allows carrying out the reading, interpretation, signal calibration, and file storage. To guaranty the signal synchronization and data coherence the acquisition and recording processes are building using the special blocks xPC Scopes, provided by the xPC-Target Toolbox. xPC Scopes are conventionally used for displaying signals in numerical or graphical form on both Target and host PCs. However, it saves defined quantity of real-time data in the RAM of the Target PC, which are accessible by the host PC through the xPC-Target API functions. This capability was herein used to obtain data uniformly sampled for displaying and recording tasks. A multithreading architecture allows acquisition processes be executed in parallel using three levels of threads. The upper level is the main thread, which is responsible for the control of xPC-Target API main loop. A second level monitors the xPC scopes states and reads the current data storage in the Target PC RAM. The last level is composed o two kind of threads, the first one (graphical) displays the real-time signal, and the second one storages all signal data necessary to storing the signals using the data persistence module.

Data Persistence Module. Data persistence was twofold accomplished. First, a relational database was implemented under the MySQL database administrator to support the user application configuration and management. Provided that the recording of biosignals require a file format that allows to add some necessary associated information, such as subject demographics, clinical diagnosis, other experimental parameters and signals, we decided to store the captured

¹ <https://github.com/java-native-access/jna>.

data using a XML-based format file, which allows to storage demographic and morphometric patient information such as binary data representing physiological signals. The content of the document changes dynamically depending on the actual study design and the anatomical structures that are examined.

3 Evaluation Setup

The purpose of this process was to evaluate the validity, reliability, and precision of the biosignals acquisition platform. Functionality the platform was twofold tested, in the former test the acquired signals were compared with reference signals provided by a function generator. In the second one, it was carried out an experimental protocol for decoding forearm and wrist movements using features extracted from electromyography and proprioceptive signals acquired using the proposed platform.

3.1 Reference Signal Comparison

In a preliminary set of tests, it was acquired sinusoidal waveforms provided by a high-accuracy calibrated digital function generator (GW Instek, model SFG-2110). A set of sinusoidal signals were generated with an amplitude of 2 V_{pp}, and following frequencies from 10 to 100 Hz with increment steps of 10 Hz, and from 100 to 500 Hz with increment steps of 100 Hz. It was checked by the follows measurements:

- Communication and quantity of data lost.
- Amplitude for acquired signals.
- Frequency distortion for acquired signals.

3.2 Decoding Wrist and Forearm Movements

Experimental Protocol. A healthy adult volunteer was instrumented with surface EMG electrodes following the SENIAM recommendations [7], provided through Ag-AgCl electrodes arranged at locations of the forearm. Two forearm muscles were measured: the wrist flexor carpi radialis and wrist extensor carpi radialis. A two-axis gyroscope (LPY510AL from STMicroelectronics) was coupled to the wrist, and another two-axis gyroscope was coupled to the forearm, which provides information of angular velocity.

A sampling frequency of 4 kHz was used to acquire records in both EMG channels and gyroscope. Within each trial, the subject repeated each limb motion three times randomly with movements: flexion and extension of the wrist, pronation, and supination of the forearm, and at resting state. The resting state was maintained at the beginning and at the end of each test, and also between movements. The order of these limb motions was randomized. A total of three trials were completed in a session.

Data Analysis. The raw EMG signals were normalized respect to the maximum voluntary contraction (MVC) value and then were full rectified by taking the absolute value of the signal. EMG signals were also filtered through a band-pass four-pole Butterworth filter, with a range of frequencies 20–500 Hz, the most widely accepted frequency limits of the EMG power spectra [8]. Signals from gyroscopes were filtered through a low-pass four-pole Butterworth filter, with a cut-off frequency of 50 Hz, taking into account that kinematic signals for human voluntary upper limb motion fall below 8 Hz [9].

Feature Extraction. Feature extraction and patterns classification algorithms were implemented in an off-line mode using functions implemented on Matlab (Mathworks Inc., Natick, MA). Frequency-domain features, as well as time-domain features, were extracted for each EMG signal and, a time-domain feature was also extracted from each proprioceptive sensor.

Time Domain Feature Extraction. Root Mean Square (RMS) was modeled as amplitude modulated Gaussian random process whose RMS was related to the constant force and non-fatiguing contraction. It relates to standard deviation, which can be expressed by the Eq. 1.

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N x_n^2} \quad (1)$$

where x_n represents the data of the segment n of N samples.

The mean absolute value (MAV) was calculated using a moving window. It was calculated for each window of data according to the Eq. 2:

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (2)$$

where x_i represents the data of the segment i of N samples length.

Frequency Domain Feature Extraction. Autoregressive (AR) model described each sample of SEMG signal as a linear combination of previous samples plus a white noise error term [10]. AR coefficients were herein used as features describing EMG pattern for the recodification. The model is basically the described by the Eq. 3:

$$X_n = \sum_{i=1}^p a_i \cdot x_{n-1} + w_n \quad (3)$$

where X_n is a sample of the model signal, a_i is an AR coefficient, w_n is white noise or error sequence, and p is the order of the AR model.

Feature Vectors. Feature vectors, which are then used in the classification stage, result from the concatenation of both time and frequency domains. Two different algorithms were evaluated (see Fig. 4).

Algorithm 1. The feature vector contains concatenated characteristics of the first four auto-regressive coefficients and the RMS value for two EMG signals. So, we have 10 features ($2 \text{ channels} \times 5 \text{ features/channel}$).

Algorithm 2. The Feature vector contains concatenated features of the first four auto-regressive coefficients and the RMS value for two EMG signals, and MAV for two proprioceptive channels. So, we have 12 features ($2 \text{ EMG channels} \times 5 \text{ features/channel} + 2 \text{ proprioceptive channels} \times 1 \text{ feature/channel}$).

Classification Stage and Post-processing. To carried out the classification process, it was used a linear classifier based on the Linear Discriminant Analysis (LDA) method [11]. In the classification process, each movement in the experimental protocol belongs to a class as following: class 1, wrist flexion; class 2, wrist extension; class 3, forearm pronation; class 4, forearm supination; class 5, resting. As three trials were completed in each session, two of them were used to training the classification model, whilst the last one was used for testing.

The post-processing methods are designed to manage excessive outflows in the classification process for improving the system performance. The majority vote method (MV) uses the current classification result, along with the n previous classification results and makes a classification decision based on the class that appears more often. The resulting effect is a smooth operation that removes spurious misclassification.

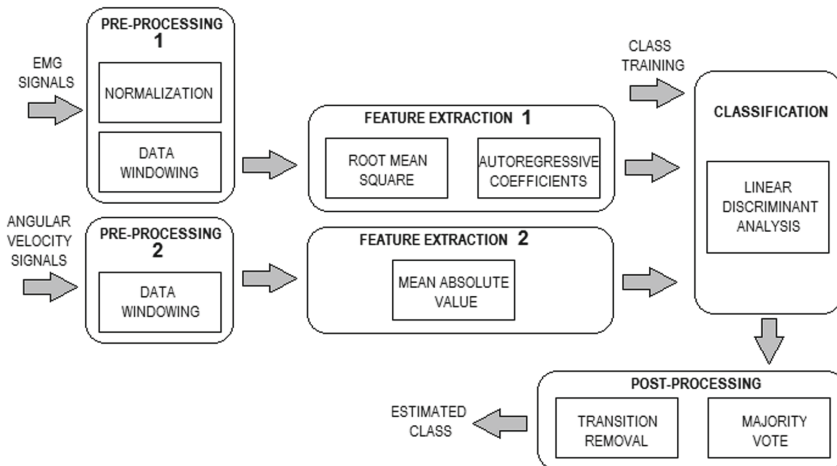


Fig. 4. Algorithm for processing time-domain and frequency-domain features from EMG signals and proprioceptive information

4 Results

4.1 Reference Signal Comparison

Communication and Quantity of Data Acquired. This test aims to verify that it does not exist loss of data. It was compared the number of samples acquired during a period of time of 20 min. It was not detected loss of information.

Measurement of Amplitude for Acquired Signals. A set of signals were acquired and stored with a sampling rate of 4000 samples per second. The root mean square error (RMSE) was used to quantify the difference from the value of the reference and the value recorded, as can be seen on Eqs. 4 and 5.

$$e(n) = \text{Amplitude}_{reference}(n) - \text{Amplitude}_{recorded}(n) \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^k e^2(n)} \quad (5)$$

The maximum error was for a signal acquired with a frequency of 500 Hz, with a RMSE of 2.67, which represents a measure of good accuracy.

Calculation of Frequency Distortion for Acquired Reference Signals.

To measure frequency distortion, it was calculate the FFT (Fast Fourier Transform) and total harmonic distortion (THD) which is defined as the ratio of the sum of the powers of all harmonic components to the power of the fundamental frequency. THD was calculated for each set of acquired signals. THD values were below of 0.5%, thus it was not found distortion for all acquired signals.

4.2 Decoding Wrist and Forearm Movements

Performance of the proposed algorithms was evaluated based on statistics measures that include mean error, sensitivity and specificity [12]. Figure 5 (left) shows the classification error of classes obtained from the confusion matrix for Algorithm 1. The main diagonal in the confusion matrix represents the concordance between the actual and the predicted class. In this case, the error of classification of wrist and forearm movements was 10.14%.

Figure 5 (right) shows the classification error of classes for Algorithm 2, using a feature vector with characteristic from EMG signals and proprioceptive information. The error of classification of wrist and forearm movements was 8.69%. Similarly, Table 1 presents the statistical values of sensitivity and specificity for each algorithm. It could be observed that Algorithm 2 presents better performance, which suggests that it is a good option for myoelectric pattern recognition, to complement EMG data with proprioceptive information.

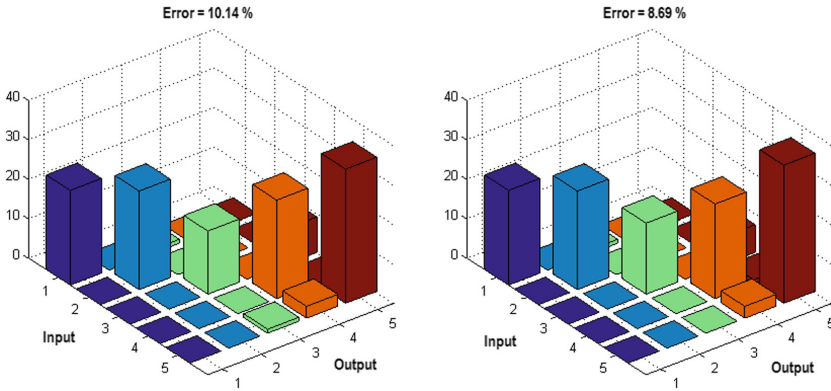


Fig. 5. Error of classification obtained from the confusion matrix. Features extracted from EMG signals (left); error of classification obtained from the confusion matrix. Features extracted from EMG signals and proprioceptive information (right)

Table 1. EMG-based pattern recognition algorithms performance

	Sensitivity (%)	Specificity (%)
Algorithm using EMG signals	89.52	82.24
Algorithm using EMG and proprioceptive data	91.07	86.77

5 Conclusions and Future Work

This work described a software platform to setup experiments for research purposes. Platform enables studies in human-machine interaction, and specifically for detecting movement intention by capturing simultaneously bioelectric and biomechanical signals at upper limb level during specific motor tasks. Developed tool can be extended to be used with other acquisition systems and processing approaches thanks to their modular architecture. Furthermore, biosignals can be recorded in real-time by a custom Java-based GUI application, and stored in memory to be latterly off-line analyzed. Thanks to these functionalities, it is possible testing, comparison and validation of pattern recognition algorithms for detecting movement intention.

For validation purposes, it was proposed an experimental protocol to compare two algorithms for myoelectric pattern recognition: an algorithm uses features only from EMG signals, and the other one uses features from EMG signals and proprioceptive information. The algorithms were validated through a set of experiments to detect five wrist and forearm. It was used a linear classifier based on LDA. Obtained results show that it was obtained an error of 10.14% with the first algorithm and it was obtained an error of 8.69% with algorithm using features EMG signals and proprioceptive information.

Acknowledgement. This work was supported by the Vice-rectory of Research of Universidad Antonio Nariño under project 2015227.

References

1. Tao, R., Xie, S., Zhang, Y., Pau, J.W.L.: Review of EMG-based neuromuscular interfaces for rehabilitation: elbow joint as an example. *Int. J. Biomech. Biomed. Robot.* **2**, 184–194 (2013)
2. Wang, L., Li, C., Wu, J.: The status of research into intention recognition. In: *Improving the Quality of Life for Dementia Patients through Progressive Detection, Treatment, and Care*, pp. 201–221. IGI Global (2017)
3. Geethanjali, P.: Myoelectric control of prosthetic hands: state-of-the-art review. *Med. Devices (Auckl. NZ)* **9**, 247 (2016)
4. Oskoei, M.A., Hu, H.: Myoelectric control systems—a survey. *Biomed. Sig. Process. Control* **2**(4), 275–294 (2007)
5. Fougner, A., Scheme, E., Chan, A.D.C., Englehart, K., Stavdahl, O.: Resolving the limb position effect in myoelectric pattern recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **19**, 644–650 (2011)
6. Mosterman, P.J., Prabhu, S., Dowd, A., Glass, J., Erkinen, T., Kluza, J., Shenoy, R.: Embedded real-time control via MATLAB, simulink, and xPC-target. In: *Handbook of Networked and Embedded Control Systems*, pp. 419–446 (2012)
7. Stegeman, D.F., Hermens, H.J.: Standards for surface electromyography: the European project surface EMG for non-invasive assessment of muscles (SENIAM). In: *Enschede: Roessingh Research and Development*, pp. 108–12 (2007)
8. Merletti, R., Parker, P.J.: *Electromyography: Physiology, Engineering, and Non-Invasive Applications*. Wiley-IEEE Press, Hoboken (2004)
9. Milner, T.E., Cloutier, C.: Damping of the wrist joint during voluntary movement. *Exp. Brain Res.* **122**, 309–317 (1998)
10. Phinyomark, A., Limsakul, C., Phukpattaranont, P.: A novel feature extraction for robust EMG pattern recognition. *Comput. Vis. Pattern Recognit.* **1**, 71–80 (2009)
11. López-Delis, A., Ruiz-Olaya, A.F., Freire-Bastos, T., Delisle-Rodríguez, D.: A comparison of myoelectric pattern recognition methods to control an upper limb active exoskeleton. In: Ruiz-Shulcloper, J., Sanniti di Baja, G. (eds.) *CIARP 2013 Part II*. LNCS, vol. 8259, pp. 100–107. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41827-3_13
12. Scheme, E., Englehart, K.: A comparison of classification based confidence metrics for use in the design of myoelectric control systems. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7278–7283. IEEE (2015)



Augmented Visualization and Touchless Interaction with Virtual Organs

Lucio Tommaso De Paolis^(✉)

AVR Lab, Department of Engineering for Innovation,
University of Salento, Lecce, Italy
lucio.depaolis@unisalento.it

Abstract. The actual trend in surgery is the transition from open procedures to minimally invasive interventions and revolutionary changes in computer-aided surgery have been obtained thanks to the introduction of the augmented reality technology that can support the doctor during the surgery. A realistic three-dimensional visualization of the patient's organs can be obtained by means of specific algorithms of segmentation and classification of medical images and advanced modalities of touchless interaction and innovative gesture-control devices permit surgeon to have a natural and simple way of fruition of the patient's data preserving the surgical environment from the danger of contamination.

Keywords: Touchless interaction · Computer aided surgery
Interactive 3D modelling · Augmented reality

1 Introduction

Minimally Invasive Surgery (MIS), such as laparoscopy or endoscopy, has changed the way to practice the surgery and the actual trend in surgery is the transition from open procedures to minimally invasive interventions that permits a visual feedback only through the laparoscope camera without a direct palpation of organs. The use of MIS reduces the surgical trauma for the patient because the diseased area is reached by means of small incisions made on the patient body and specific instruments and a camera are inserted through these ports. The surgeon uses a monitor to see what happens on the surgical field inside the patient body.

The Augmented Reality (AR) technology has the potential to improve the surgeon's visualization during the MIS surgery. This technology can "augment" surgeon's perception of the real world with the use of information gathered from patient's medical images. AR technology refers to a perception of a physical real environment whose elements are merged with virtual computer-generated objects in order to create a mixed reality. The merging of virtual and real objects in an AR application has to run in real time.

The use of AR technology in medicine makes possible to overlay virtual medical images of the organs on the real patient. This allows the surgeon to have a sort of "X-ray vision" of the patient's internal anatomy. The use of AR in surgery produces a better spatial perception and a reduction in the duration of the surgical procedure.

In order to obtain a correct alignment of virtual and real objects is needed a registration procedure that can be accomplished using an optical tracker [1].

Medical images (CT or MRI) associated to the latest medical image processing techniques could provide an accurate knowledge of the patient's anatomy and pathologies. This information is used in augmented visualization in order to guide surgeons during the surgical procedure and to improve the patient care [2].

Many research groups are now focusing on the development of systems that assists surgeons during the minimally invasive surgical procedures using the augmented reality technology.

Samset et al. [3] present some decision support tools based on concepts in visualization, robotics and haptics and provide tailored solutions for a range of clinical applications.

Bichlmeier et al. [4] focus on the problem of misleading perception of depth and spatial layout in medical AR and present a new method for medical in-situ visualization.

Nicolau et al. [5] present a real-time superimposition of virtual models over the patient's abdomen in order to have a three dimensional view of the internal anatomy. The authors have used the developed system in an operating room and to reduce the influence of the liver breathing motion they have tried to simulate the patient's breathing cycle.

De Mauro et al. [6] present the development of an IGS solution for spine surgery based on conventional medical devices and open-source software. The aim of the work is to enhance the surgeon's ability for a better intra-operative orientation by giving a three-dimensional view and other information necessary for a safe navigation inside the patient. The system allows reducing the radiation exposure to the patient and medical staff as well as the operation duration.

De Paolis et al. [7] present an application based on the AR Toolkit developed in order to visualize all the organs in a given abdominal area or only some of them. The system can be used in pre-operative planning of a laparoscopic surgical procedure to find both the right points to introduce the surgical instruments and the best path to reach the organ concerned in the surgical operation.

Ricciardi et al. [8] present an augmented reality platform for computer-assisted surgery in the field of maxillo-facial surgery. A surgical planning module is also integrated in the platform.

De Paolis et al. [9] present a guidance system based on AR for the needle placement in the treatment of the liver tumours with the radiofrequency ablation. The software offers also a set of tools to study the patient's pathology during the preoperative planning phase.

In orthopaedic surgery, AR has been used in image-guided procedures such as percutaneous vertebroplasty. Using a head-mount display and a marker sheet placed on the patient, the virtual images have allowed showing the surgeon the correct angle of needle insertion. Results from this study showed a decreased error when inserting the needle was done with the AR-guided procedures compared to cases in which it was not used [10].

Baker et al. [11] present the results of a recent application involving AR to communicate with remote surgeons during surgery. The local surgeon wears the Google Glass and the remote surgeon projects his hands into the local surgeon's viewing field. The authors underline that experienced surgeons may virtually provide real-time

assistance to a less experienced surgeon and that the AR can be used in surgical simulation settings to improve the resident training.

2 Building of the 3D Model of the Organs

One of the biggest limits of most traditional medical imaging techniques is the bi-dimensionality of the images. The radiologist and/or the surgeon need a training phase in order to correctly interpret the content of medical images, to map grey levels in the images to the different kind of anatomical tissues and to reconstruct the three-dimensional structure from the two-dimensional sections of the organs.

For this reason medical imaging is moving towards the creation of three-dimensional models of the patient organs not only for a better comprehension of the patient anatomy, but also to improve the communication between physician and patient [12, 13]. A stereo viewing of the internal anatomy allows a strong improvement in users' performance and leads to an accurate navigation and a faster decision-making from the surgeon [14, 15].

An efficient three-dimensional reconstruction of the patient's organs is generated by applying to the medical images (CT or MRI) some segmentation and classification algorithms and the grey levels of the medical images are replaced with a colour map associated to different organs [16].

Image segmentation usually represents the core of image analysis and it consists in the partitioning of a dataset in contiguous regions (anatomical and pathologic structures) that have common properties. Usually, the results of the segmentation algorithm are enhanced by some kind of post-processing, such as the smoothing of the boundary or the removal of small holes.

In Fig. 1 is shown the 3-dimensional model of the heart.

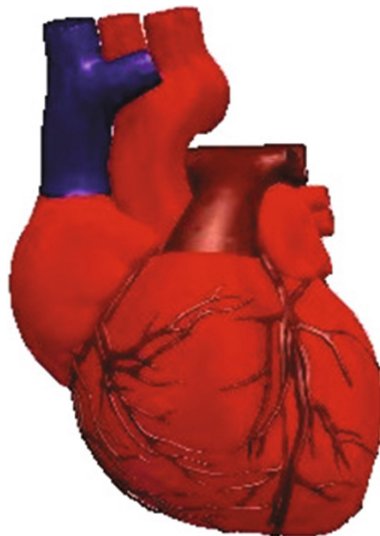


Fig. 1. Three-dimensional model of the heart.

3 Gestural Technologies

Gestures are new forms of communication based on the association of particular messages and meaningful commands with well-defined positions or movements of some parts of the human body. They typically deal with finger and hand movements [17]. Gestures represent a more natural form of human-computer interaction than traditional devices such as mouse and keyboard. In gesture-based environments, users can pay more attention to the output on the screen, since they do not have to look at the input devices anymore. Moreover, gestures are a more practical alternative to voice-recognition systems, which require a long and complex training phase to adjust to the voice tone and the users diction.

In the last years a lot of devices and systems have been designed for gesture detection. However, hints provided by human-interaction experts should be taken into account in the design process to make gestural technologies more usable and consistent with natural interaction forms. Moreover, the distinction between real control gestures and accidental movements is still an open problem.

In the earlier interaction systems based on hand gestures users had to wear electronic gloves containing several sensors. An important drawback coming from the adoption of electronic gloves is related to the calibration of those systems, whose parameters need to be tuned according to hand geometry [18].

Some other gesture detection systems apply color segmentation on colored gloves worn by users [19] or use vision-tracked fiducial markers placed on some handheld devices or even directly over hands and fingertips [20].

Wearable devices are often cumbersome and may limit hand movements due to the presence of sensors and wires. The consequent constraints on the degree of freedom of movements partially reduce the range of users' gestures. Furthermore, the distinction between real control gestures and accidental movements is still an open problem.

Generally, the surgeon needs to consult during surgery some digitalized data related to the patient's medical records that are displayed on a monitor placed in the operating room. The main limitation, in such a situation, is that he handles surgical instruments and, above all, that he wears sterile gloves and coat and he cannot in any way come into contact with IT equipment that are potentially vehicle of bacteria.

Thus, as often happens, in this case the surgeon has to avail the help of medical staff at the expense of precision and time optimization. In this perspective, the goal of the developed system would be to enhance the practice of the surgeon who, by using an innovative wearable device that allows touchless gesture-control, can both interact with 3D models of the patient's organs and consult medical information of the traditional slices in an absolutely autonomous way, simply by performing natural arm and hand movements.

Nowadays the most common gestural technologies, introduced by videogame and entertainment companies (e.g., Microsoft, Nintendo, and Sony), rely either on handheld devices or on cameras performing motion tracking [21].

The Nintendo Wii remote controller (Wiimote) [22] is equipped with accelerometers, infrared detectors, and a LED sensor bar.

The Leap Motion controller [23] is a small, easy to-use, and low-cost device designed to capture the movements of human hands and fingers. Furthermore, the Leap Motion controller allows a more natural form of interaction: users do not have to hold any particular object and, thus, can freely move their hands and fingers in the space.

The Sony PlayStation Move controller [24] detects movements in the 3D space thanks to a multi-colored light source and a webcam.

Microsoft introduced Kinect [25], a markerless webcam-based system that does not need any handheld device.

Myo Armband is a wearable gesture-control and motion-control device developed by Thalmic Labs [26]; it is shown in Fig. 2.



Fig. 2. The Myo Armband

Myo it's a device that employs the extraordinary capability of eight separate electromyography (EMG) metal sensors that, in contact with the skin, are able to literally read the electrical activity of forearm muscles and to translate them into digital commands. From here its name "Myo", the scientific prefix for muscles. Other sensors, such as compass, gyroscope and accelerometer, can identify the arm on which it was worn and detect its rotations and orientation. The Myo Armband working modality is shown in Fig. 3.

The Bluetooth communication technology provides surgeon with different benefits:

- Myo can be worn under the white coat thus preserving the surgical environment from external contamination;
- the surgeon doesn't need to point the arm in any specific direction, thus feeling free to assume the more comfortable position.

The preconfigured 5 gestures (double tap, fist, fingers spread, wave left and wave right) could be combined with forearm movements and twists as shown in Fig. 4.

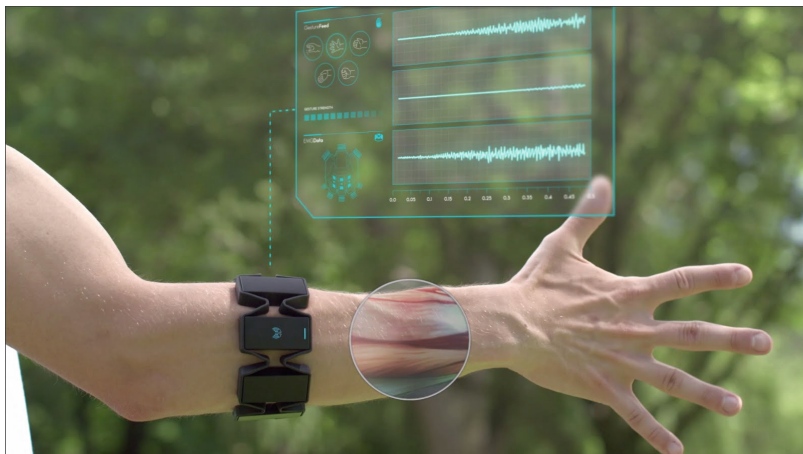


Fig. 3. The Myo Armband working modality.

Wave Left			Leftwards Rotation
Wave Right			Rightwards Rotation
Spread fingers			Zoom out
Fist			Zoom in
Double tap			Interaction Lock

Fig. 4. The 5 preconfigured gestures of Myo Armband.

4 Interaction with Virtual Organs

With the aim to design the touchless control of the application and implement its interactive behaviour, the “Keyboard Mapper” utility provided with the Myo Connect has been used. Each of the 5 gestures that Myo recognizes has been associated with one of the specific keyboard commands and with an additional very important task.

The association between gestures the surgeon performs and commands executed on the virtual models has not been made in a random manner but in such a way to provide the surgeon an experience of interaction as natural as possible.

So, assuming the Myo worn on the right arm, the resulting mapping has been the following:

- Wave Left - left arrow key: by moving the hand towards the inside, surgeon can intuitively rotate the virtual object leftwards
- Wave Right - right arrow key: similarly, by moving the hand towards the outside, surgeon can easily rotate the virtual object rightwards;
- Fist - down arrow key: surgeon can zooming in the augmented element by closing the hand in a fist, as if he wanted grab it;
- Spread Fingers - up arrow key: surgeon can zooming out the augmented element by spreading his fingers, as if he wanted move it away;
- Double Tap - toggle lock: to turn on and off the motion control system.

The last mapping provides surgeon with very important additional gesture. It deals with the lock of the device that allows the surgeon in any time to interrupt the touchless interaction and carries out the surgery by prevent the interpretation of accidental gestures.

In Fig. 5 is shown the interaction with virtual organs using the Myo Armband.



Fig. 5. The interaction with virtual organs using the Myo Armband.

5 Interaction with Medical Images

In order to make surgeon able to hands-free consult also the slices of the medical images displayed on a monitor, the default “Presentation Mode” built in the Myo Connect utility has been used. It can be activated through the “Myo Application Manager” window just flipping the switch on and allows surgeon to interact with the slices by performing just three simple gestures:

- Double Tap: move forward in the presentation
- Wave Left and Hold: holding the wave left gesture until Myo stops vibrating to move backward in the presentation
- Make a Fist and Rotate it's a combination of gestures that activates:
 - Pointer: if the fist is rotated counter-clockwise; Double Tap to turn the pointer off;
 - Zoom-in: if the fist is rotated clockwise; Double Tap to zoom out.

Finally, thanks to the Bluetooth technology used for the connection with the computer system, this device is really useful in this specific use case. In fact, as long as it is in contact with the skin, the surgeon could wear it under the white coat without compromising the communication signal and thus preserving the surgical environment from a possible external contamination.

In addition, his arm doesn't have to be pointed in any direction so he feels free to assume the more comfortable position.

In Fig. 6 is shown the interaction with medical images using the Myo Armband.



Fig. 6. The interaction with medical images using the Myo Armband.

6 Conclusions and Future Work

The goal of the developed system is to permit surgeon by using the Mayo Armband, an innovative wearable device that allows touchless gesture-control, to interact with 3D models of the patient's organs and consult medical traditional slices in an absolutely autonomous way simply by performing natural arm and hand movements. Myo it's a

device that employs the capability of some electromyography metal sensors that, in contact with the skin, are able to literally read the electrical activity of forearm muscles and to translate them into digital commands.

The results of this work look extremely interesting and led to the development of a first prototype of an application very useful, easy to use and that does not completely upset the “modus operandi” already consolidated in the surgeon’s experience.

The association between gestures and tasks has been made in such way to provide the surgeon an experience of interaction as natural as possible. The system guarantees absolute autonomy in consulting the patient’s medical data, a simple touchless interaction through natural gestures of the hand and arm, a considerable decrease of the time of access to the patient’s data.

With a view to the future development of this work, the objective is to make the rendering of both graphic aspect and real natural movements of the organs (such as respiratory movement or heartbeat) to be synergically integrated in real time and not virtually reconstructed.

Then, it will be the possibility to intervene in greater detail on the libraries that implement the functioning of the Myo device in order to implement the recognition of additional gestures in order to provide the surgeon the ability to interact with the virtual models of the patient’s organs by exploiting a wider range of operations.

Finally, it will certainly be subject of future study also the possibility to realize the rendering of the virtual organs with advanced AR techniques which aim to display them in the form of holograms. In this way, the surgeon would have the ability to interact with them using an extra degree of freedom.

References

1. Maad, S.: Augmented reality. In: *The Horizon of Virtual and Augmented Reality: The Reality of the Global Digital Age*. Intech, January 2010. ISBN 978-953-7619-69-5
2. De Paolis, L.T., Aloisio, G.: Augmented reality in minimally invasive surgery. In: Mukhopadhyay, S.C., Lay-Ekuakille, A. (eds.) *Advances in Biomedical Sensing, Measurements*. LNEE, vol. 55. Instrumentation and Systems. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-05167-8_17
3. Samset, E., Schmalstieg, D., Vander Sloten, J., Freudenthal, A., Declerck, J., Casciaro, S., Rideng, Ø., Gersak, B.: Augmented reality in surgical procedures. In: *SPIE Human Vision and Electronic Imaging XIII* (2008)
4. Bichlmeier, C., Wimmer, F., Michael, H.S., Nassir, N.: Contextual anatomic mimesis: hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: *Proceedings of Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2007)*, Nara, Japan, pp. 129–138 (2007)
5. Nicolau, S.A., Pennec, X., Soler, L., Buy, X., Gangi, A., Ayache, N., Marescaux, J.: An augmented reality system for liver thermal ablation: design and evaluation on clinical cases. *Med. Image Anal.* **13**(3), 494–506 (2009). Elsevier
6. De Mauro, A., Mazars, J., Manco, L., Mataj, T., Fernandez, A.H., Cortes, C., De Paolis, L. T.: Intraoperative navigation system for image guided surgery. In: *6th International Conference on Complex, Intelligent, and Software Intensive Systems*, pp. 486–490 (2012)

7. De Paolis, L.T., Pulimeno, M., Aloisio, G.: An augmented reality application for minimally invasive surgery. In: Katashev, A., Dekhtyar, Y., Spigulis, J. (eds.) 14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics. IFMBE, vol. 20, pp. 489–492. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-69367-3_131
8. Ricciardi, F., Copelli, C., De Paolis, L.T.: A pre-operative planning module for an augmented reality application in maxillo-facial surgery. In: De Paolis, L.T., Mongelli, A. (eds.) AVR 2015. LNCS, vol. 9254, pp. 244–254. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22888-4_18
9. De Paolis, L.T., Ricciardi, F., Dragoni, A.F., Aloisio, G.: An augmented reality application for the radio frequency ablation of the liver tumors. In: Murgante, B., Gervasi, O., Iglesias, A., Taniar, D., Apduhan, B.O. (eds.) ICCSA 2011. LNCS, vol. 6785, pp. 572–581. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21898-9_47
10. Abe, Y., Sato, S., Kato, K., Hyakumachi, T., Yanagibashi, Y., Ito, M., Abumi, K.: A novel 3D guidance system using augmented reality for percutaneous vertebroplasty. *J. Neurosurg. Spine* **19**, 492–501 (2013)
11. Baker, D.K., Fryberger, C.T., Ponce, B.A.: The emergence of augmented reality in orthopaedic surgery and education. *Orthop. J. Harvard Med. School* **16**, 8–16 (2015)
12. Houts, P.S., Doak, C.C., Doak, L.G., Loscalzo, M.J.: The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Educ. Counsel* **61**, 173–190 (2006)
13. König, T.S.: Usability issues in 3D medical visualization. Ph.D thesis, Technische Universität Wien, Wien (2001)
14. Livatino, S., De Paolis, L.T., D’Agostino, M., Zocco, A., Agrimi, A., De Santis, A., Bruno, L.V., Lapresa, M.: Stereoscopic visualization and 3-D technologies in medical endoscopic teleoperation. *IEEE Trans. Ind. Electron.* **62**(1), 525–535 (2015)
15. De Paolis, L.T., De Mauro, A., Raczkowski, J., Aloisio, G.: Virtual model of the human brain for neurosurgical simulation. *Stud. Health Technol. Inform.* **150**, 811–815 (2009)
16. Yoo, T.S.: *Insight into Images*. A K Peters, Natick (2004)
17. Garber, L.: Gestural technology: moving interfaces in a new direction [technology news]. *Computer* **46**, 22–25 (2013)
18. Joslin, C., El-Sawah, A., Chen, Q., Georganas, N.: Dynamic gesture recognition. In: Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Ottawa, ON, Canada, 16–19 May 2005, vol. 3, pp. 1706–1711 (2005)
19. Bellarbi, A., Benbelkacem, S., Zenati-Henda, N., Belhocine, M.: Hand gesture interaction using color-based method for tabletop interfaces. In: Proceedings of the 7th International Symposium on Intelligent Signal Processing (WISP), Floriana, Malta, 19–21 September 2011, pp. 1–6 (2011)
20. Fiala, M.: ARTag, a fiducial marker system using digital techniques. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, San Diego, CA, USA, 20–25 June 2005, vol. 2, pp. 590–596 (2005)
21. Invitto, S., Faggiano, C., Sammarco, S., De Luca, V., De Paolis, L.T.: Haptic, virtual interaction and motor imagery: entertainment tools and psychophysiological testing. *Sensors* **16**, 394 (2016)
22. Nintendo Wii System. <http://wii.com>
23. The leap motion controller. <http://www.leapmotion.com>
24. Sony PlayStation. <http://www.playstation.com>
25. Microsoft Kinect. <http://support.xbox.com/en-US/browse/xbox-one/kinect>
26. Myo Armband. <http://www.myo.com>



Decreased Composite Indices of Femoral Neck Strength in Young Obese Women

Abdel-Jalil Berro^{1,2}, Said Ahmaidi², Antonio Pinti⁸,
Abir Alwan^{1,3}, Hayman Saddik^{1,8}, Joseph Matta^{3,4,5,6},
Fabienne Frenn^{3,4,5}, Maroun Rizkallah^{3,4,7}, Ghassan Maalouf⁷,
and Rawad El Hage^{1(✉)}

- ¹ Department of Physical Education, Faculty of Arts and Social Sciences,
University of Balamand, El-Koura, Lebanon
rawadelhage21@hotmail.com
- ² EA-3300, APERE, Sport Sciences Department,
University of Picardie Jules Verne, Amiens, France
- ³ Movement, Sport and Health Sciences Laboratory (M2S), UFR-APS,
University of Rennes 2-ENS Cachan, Rennes, France
- ⁴ Laboratoire VIP'S, UFR-APS, Campus la Harpe,
Université Rennes 2, Rennes, France
- ⁵ Industrial Research Institute, Beirut, Lebanon
- ⁶ Faculty of Pharmacy, Department of Nutrition,
Saint Joseph University, Beirut, Lebanon
- ⁷ Faculty of Medicine, Bellevue University Medical Center,
Saint Joseph University, Mansourieh, Lebanon
- ⁸ I3MTO, University of Orléans, Orléans, France

Abstract. The aim of the current study was to compare compression strength index (CSI), bending strength index (BSI) and impact strength index (ISI) among obese, overweight and normal-weight young women. 117 young women (20 obese, 36 overweight and 61 normal-weight) whose ages range from 18 to 35 years participated in this study. Body composition and BMD were evaluated by dual-energy X-ray absorptiometry (DXA). CSI, BSI and ISI values were significantly lower in obese and overweight women compared to normal-weight women ($p < 0.001$). In the whole population ($n = 117$), body mass index (BMI) was negatively correlated to CSI ($r = -0.66$; $p < 0.001$), BSI ($r = -0.56$; $p < 0.001$) and ISI ($r = -0.54$; $p < 0.001$). This study suggests that obesity is associated with lower CSI, BSI and ISI values in young women.

Keywords: Bone strength · Fracture risk · Hip geometry

1 Introduction

Recent studies have indicated that obesity may increase fracture risk at the ankle, the lower limbs, the humerus and the spine [1, 2]. There are several factors that may explain the development of bone fragility in obese women [1, 2]. These factors include low serum levels of vitamin D, high serum parathyroid levels, low serum testosterone levels, increased production of pro-inflammatory and pre – resorptive cytokines by the

visceral fat tissue [1–3]. Moreover, obese subjects seem to have low BMD values for their body weight [3–5]. BMD is generally considered as the best determinant of bone strength and can therefore practically predict fracture risk [6]. However, only 50 to 70% of bone strength variability can be explained by BMD [6, 7]. Hip bone strength is also influenced by other factors such as femoral neck width and bending strength [8]. Karlamangla et al. [8] have examined the prediction of incident hip fracture risk by composite indices of femoral neck strength (Compression strength index, bending strength index and impact strength index) constructed from DXA scans of the hip. These indices integrate femoral neck size and body size with bone density [8–10]. Compression strength index (CSI), bending strength index (BSI) and impact strength index (ISI) reflect the ability of the femoral neck to withstand axial compressive and bending forces and the ability to absorb energy from an impact [8–10]. These indices have been shown to improve hip fracture risk and bone strength assessments in the elderly [8–10]. We have recently shown that obesity is associated with lower CSI, BSI and ISI values in young men [11]. However, little is known concerning the relation between obesity and composite indices of femoral neck strength in young obese women. The aim of the current study was to compare CSI, BSI and ISI among obese, overweight and normal-weight young women.

2 Methods

117 young women (20 obese, 36 overweight and 61 normal-weight) whose ages range from 18 to 35 years participated in this study. Normal-weight, overweight, and obesity are defined by a BMI between 20 and 24.9 kg/m², between 25 and 29.9 kg/m², and 30 kg/m² or more, respectively. The participants were recruited from 4 private universities located in North Lebanon. All participants were non - smokers and had no history of major orthopedic problems or other disorders known to affect bone metabolism or physical tests of the study. Other inclusion criteria included no diagnosis of comorbidities and no history of fracture. An informed written consent was obtained from the participants. This study was approved by the University of Balamand Ethics Committee. Bone mineral content (BMC) and BMD were determined for each individual by DXA at whole body (WB), lumbar spine (L1-L4), total hip (TH), and femoral neck (FN) (GE Healthcare, Madison, WI). Body composition was also evaluated by DXA. Composite indices of femoral neck strength (CSI, BSI and ISI) were calculated as previously described [8–10].

The ability of the femoral neck to withstand compressive and bending forces proportional to body weight can be approximated by the following indices:

Compression strength index = $(\text{BMD} * \text{Femoral neck width}) / \text{body weight}$,
and

Bending strength index = $(\text{BMD} * \text{Femoral neck width}^2) / (\text{hip axis length} * \text{body weight})$.

The ability of the femoral neck to absorb the energy of impact in a fall from standing height is given by:

Impact strength index = (BMD * Femoral neck width * hip axis length)/(body height * body weight).

Basic data are presented as mean \pm standard deviation (SD). Comparisons between the three groups were made after checking for Gaussian distribution. If Gaussian distribution was found, one-way analysis of variance was used. In the other case, one-way analysis of variance on ranks test was used. Associations between clinical characteristics and bone data were given as Pearson correlation coefficients. The difference was considered statistically significant at $p < 0.05$. Data were analyzed using NCSS (2001).

3 Results

Age and height were not significantly different among the three groups. Weight, BMI, lean mass, fat mass and fat mass percentage were significantly higher in obese and overweight women compared to normal-weight women ($p < 0.01$). WB BMC and WB BMD were significantly higher in obese women compared to normal-weight women ($p < 0.05$). Lumbar spine BMD was not significantly different among the three groups. TH BMD and FN BMD were significantly higher in obese women compared to normal-weight women ($p < 0.001$). CSI, BSI and ISI values were significantly lower in obese and overweight women compared to normal-weight women ($p < 0.001$). In the whole population ($n = 117$), BMI was negatively correlated to CSI ($r = -0.66$; $p < 0.001$), BSI ($r = -0.56$; $p < 0.001$) and ISI ($r = -0.54$; $p < 0.001$) but positively

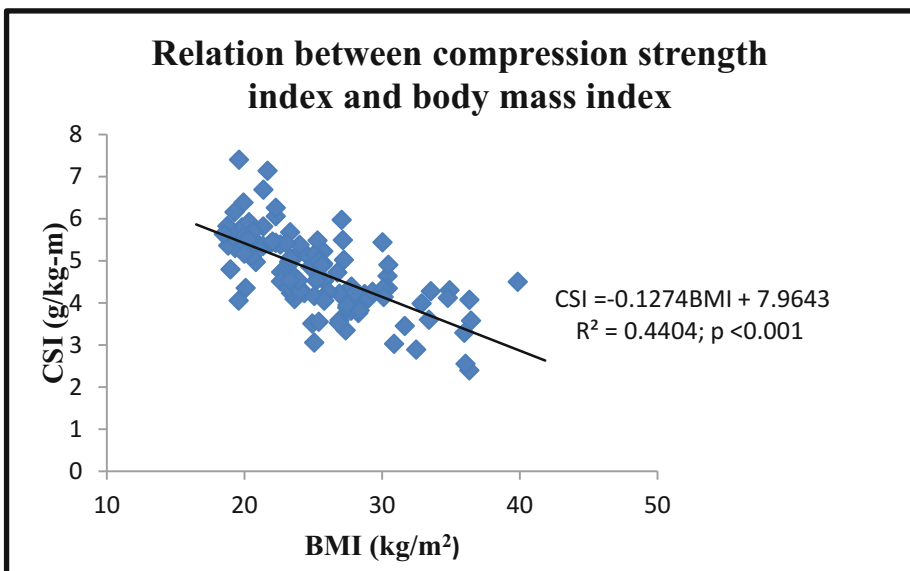


Fig. 1. Relation between CSI and BMI in the whole population. R² is the proportion of the variance in the dependent variable that is predictable from the independent variable; p -value: probability value.

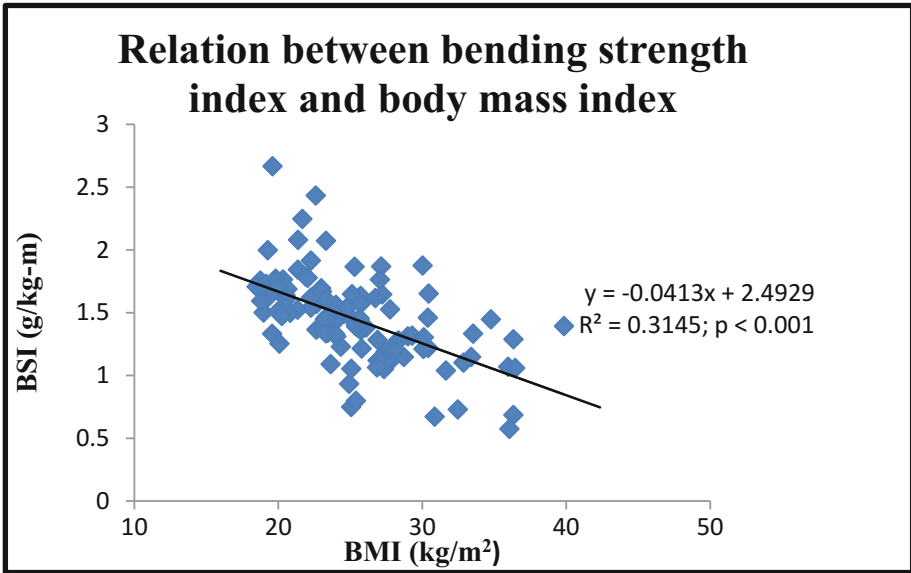


Fig. 2. Relation between BSI and BMI in the whole population. R^2 is the proportion of the variance in the dependent variable that is predictable from the independent variable; p -value: probability value.

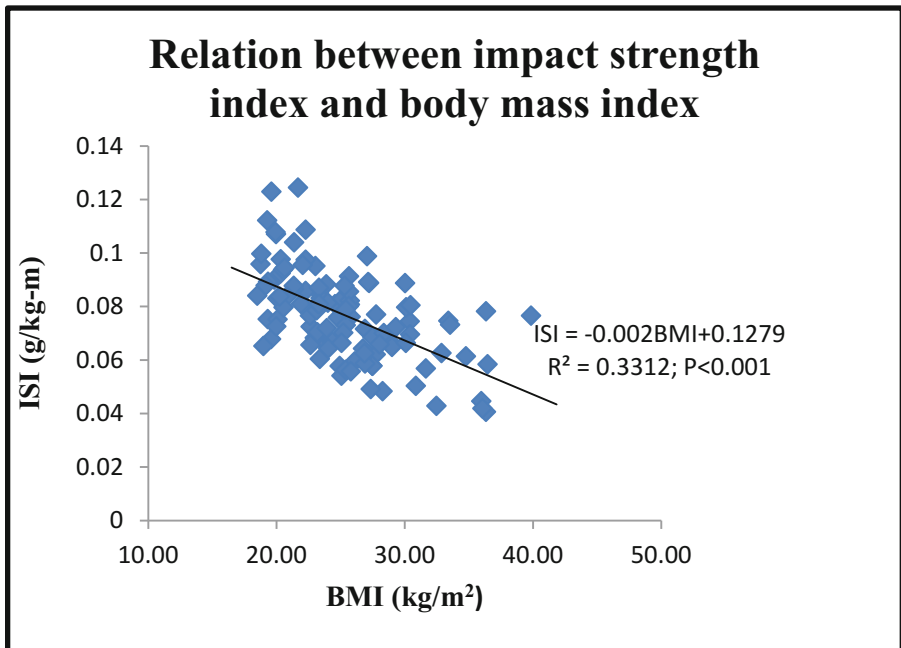


Fig. 3. Relation between ISI and BMI in the whole population. R^2 is the proportion of the variance in the dependent variable that is predictable from the independent variable; p -value: probability value.

correlated to WB BMC ($r = 0.27$; $p < 0.01$), WB BMD ($r = 0.36$; $p < 0.001$), TH BMD ($r = 0.32$; $p < 0.001$) and FN BMD ($r = 0.27$; $p < 0.01$). Moreover, body weight was negatively correlated to CSI ($r = -0.58$; $p < 0.001$), BSI ($r = -0.50$; $p < 0.001$), ISI ($r = -0.38$; $p < 0.001$) but positively correlated to WB BMC ($r = 0.45$; $p < 0.001$), WB BMD ($r = 0.47$; $p < 0.001$), TH BMD ($r = 0.38$; $p < 0.001$) and FN BMD ($r = 0.39$; $p < 0.001$). Finally, fat mass percentage was negatively correlated to CSI ($r = -0.62$; $p < 0.001$), BSI ($r = -0.52$; $p < 0.001$), ISI ($r = -0.46$; $p < 0.001$) but positively correlated to WB BMD ($r = 0.20$; $p < 0.05$). (Figs. 1, 2, 3 and Tables 1, 2)

Table 1. Clinical characteristics and bone variables in the whole population

	Normal-weight women (n = 62)	Overweight women (n = 36)	Obese women (n = 20)	p-value
Age (years)	22.4 ± 2.4	24.4 ± 3.9	22.2 ± 3.0	p = 0.072
Weight (Kg)	55.3 ± 5.2*	67.2 ± 6.6 @	85.4 ± 11.3 \$	p < 0.001
Height (m)	1.60 ± 0.058	1.58 ± 0.07	1.59 ± 0.07	p = 0.240
BMI (kg/m ²)	21.4 ± 1.8*	26.7 ± 1.2 @	33.3 ± 2.8 \$	p < 0.05
FM (kg)	18.215 ± 3.649*	28.001 ± 4.21@	39.489 ± 7.067 \$	p < 0.05
FM (%)	32.7 ± 4.7*	41.5 ± 3.9 @	46.0 ± 3.5 \$	p < 0.001
LM (Kg)	34.616 ± 56.769*	36.609 ± 6.6 @	44.671 ± 5.093 \$	p < 0.05
WB BMD (g/cm ²)	1.056 ± 0.0920	1.080 ± 0.077@	1.153 ± 0.0954 \$	p < 0.001
WB BMC (g)	2078 ± 376	2210 ± 253	2351 ± 367 \$	p < 0.05
L1-L4 BMD (g/cm ²)	1.137 ± 0.136	1.129 ± 0.132	1.217 ± 0.150	p = 0.05
FN BMD (g/cm ²)	0.942 ± 0.126	0.95 ± 0.131 @	1.056 ± 0.139 \$	p < 0.05
TH BMD (g/cm ²)	0.962 ± 0.115	0.97 ± 0.107 @	1.082 ± 0.134 \$	p < 0.001
CSI (g/kg-m)	5.263 ± 0.755*	4.37 ± 0.664 @	3.893 ± 0.786 \$	p < 0.001
BSI (g/kg-m)	1.620 ± 0.304*	1.334 ± 0.262	1.171 ± 0.340 \$	p < 0.001
ISI (g/kg-m)	0.085 ± 0.015*	0.069 ± 0.0163	0.064 ± 0.0148 \$	p < 0.001

BMI: Body Mass Index; FM: Fat Mass; BMC: Bone Mineral Content; BMD: Bone Mineral Density; WB: Whole Body; TH: Total Hip; FN: Femoral Neck; CSI: Compression Strength Index; BSI: Bending Strength Index; ISI: Impact Strength Index; *significant differences between normal-weight and overweight women, $p < 0.05$; \$ significant differences between normal-weight and obese women, $p < 0.05$; @ significant differences between overweight and obese women, $p < 0.05$.

Table 2. Correlations between clinical characteristics and bone variables in the whole population

	Weight (Kg)	Height (m)	BMI (kg/m ²)	TFM (kg)	FM (%)	LM (Kg)
WB BMD (g/cm ²)	0.47***	0.296**	0.361***	0.377***	0.205*	0.380***
WB BMC (g)	0.45***	0.448***	0.273**	0.349***	0.155	0.645***
L1-L4 BMD (g/cm ²)	0.23*	0.259**	0.140	0.129	-0.0203	0.217*
FN BMD (g/cm ²)	0.39***	0.307***	0.279**	0.320***	0.172	0.341***
TH BMD (g/cm ²)	0.38***	0.190*	0.322***	0.321***	0.173	0.346***
CSI (g/kg-m)	-0.58***	0.0813	-0.664***	-0.628***	-0.63***	-0.242*
BSI (g/kg-m)	-0.50***	0.0499	-0.561***	-0.526***	-0.52***	-0.189*
ISI (g/kg-m)	-0.38***	0.321***	-0.546***	-0.463***	-0.46***	-0.131

Note: The results shown represent r values. BMI: Body Mass Index; TFM: Total Fat Mass; BMC: Bone Mineral Content; BMD: Bone Mineral Density; WB: Whole Body; TH: Total Hip; FN: Femoral Neck; CSI: Compression Strength Index; BSI: Bending Strength Index; ISI: Impact Strength Index; *** p < 0.001; ** p < 0.01; * p < 0.05.

4 Conclusion

In our study, body weight and BMI were positively correlated to total hip and femoral neck BMD. Obese women had higher FN BMD values compared to normal-weight women. However, this rise in BMD may not be adapted to the increased body weight in obese women. Therefore, the current study suggests that obesity is associated with lower CSI, BSI and ISI values in young women. Implementing strategies to increase femoral neck strength indices in young obese women may be useful for preventing osteoporotic fractures later in life.





Conflicts of Interest. The authors state that they have no conflicts of interest.

References

1. Compston, J.E., Watts, N.B., Chapurlat, R., et al.: Obesity is not protective against fracture in postmenopausal women: GLOW. *Am. J. Med.* **124**, 1043–1050 (2011)
2. Spaine, L.A., Bollen, S.R.: ‘The bigger they come...’: the relationship between body mass index and severity of ankle fractures. *Injury* **27**, 687–689 (1996)
3. Goulding, A., Taylor, R.W., Jones, I.E., et al.: Overweight and obese children have low bone mass and area for their weight. *Int. J. Obes. Relat. Metab. Disord.* **24**, 627–632 (2000)
4. Rocher, E., Chappard, C., Jaffre, C., et al.: Bone mineral density in prepubertal obese and control children: relation to body weight, lean mass, and fat mass. *J. Bone Miner. Metab.* **26**, 73–78 (2008)
5. Rocher, E., El Hage, R., Chappard, C., et al.: Bone mineral density, hip bone geometry, and calcaneus trabecular bone texture in obese and normal-weight children. *J. Clin. Densitom.* **16**, 244–249 (2013)
6. Kanis, J.A., McCloskey, E.V., Johansson, H., et al.: A reference standard for the description of osteoporosis. *Bone* **42**, 467–475 (2008)
7. Stone, K.L., Seeley, D.G., Lui, L.Y., et al.: BMD at multiple sites and risk of fracture of multiple types: long-term results from the study of osteoporotic fractures. *J. Bone Miner. Res.* **18**, 1947–1954 (2003)
8. Karlamangla, A.S., Barrett-Connor, E., Young, J., et al.: Hip fracture risk assessment using composite indices of femoral neck strength: the Rancho Bernardo study. *Osteoporos. Int.* **15**, 62–70 (2004)
9. Ayoub, M.L., Maalouf, G., Bachour, F., et al.: DXA-based variables and osteoporotic fractures in Lebanese postmenopausal women. *Orthop. Traumatol. Surg. Res.* **100**, 855–858 (2014)
10. Yu, N., Liu, Y.J., Pei, Y., et al.: Evaluation of compressive strength index of the femoral neck in Caucasians and Chinese. *Calcif. Tissue Int.* **87**, 324–332 (2010)
11. El Khoury, C., Toumi, H., Lespessailles, et al.: Decreased composite indices of femoral neck strength in young obese men. *J. Clin. Densitom.* **20**, 268–270 (2017)



On the Use of Decision Trees Based on Diagnosis and Drug Codes for Analyzing Chronic Patients

Cristina Soguero-Ruiz¹  , Ana Alberca Díaz-Plaza¹, Pablo de Miguel Bohoyo² , Javier Ramos-López¹, Manuel Rubio-Sánchez¹, Alberto Sánchez¹, and Inmaculada Mora-Jiménez¹ 

¹ Rey Juan Carlos University, Madrid, Spain
{cristina.soguero,javier.ramos,manuel.rubio,alberto.sanchez,immaculada.mora}@urjc.es, a.alberca@alumnos.urjc.es

² University Hospital of Fuenlabrada, Madrid, Spain
pablo.miguel@salud.madrid.org

Abstract. Diabetes mellitus (DM) and essential hypertension (EH) are chronic diseases more prevalent every year, both independently and jointly. To gain insights about the particularities of these chronic conditions, we study the use of decision trees as a tool for selecting discriminative features and making predictive analyses of the health status of this kind of chronic patients. We considered gender, age, ICD9 codes for diagnosis and ATC codes for drugs associated with the diabetic and/or hypertensive population linked to the University Hospital of Fuenlabrada (Madrid, Spain) during 2012. Results show a relationship among DM/EH and diseases/drugs related to the respiratory system, mental disorders, or the musculoskeletal system. We conclude that drugs are quite informative, collecting information about the disease when the diagnosis code is not registered. Regarding predictive analyses, when discriminating patients with EH-DM and just one of these chronic conditions, better accuracy is obtained for EH (85.4%) versus DM (80.1%).

Keywords: Chronic health status · Diabetes · Hypertension
Feature selection · Decision trees · Diagnoses · Drugs

1 Introduction

Chronic conditions are becoming problematic in developed regions due to increased life expectancy. This demographic transition has altered rankings of the most predominant illnesses, where chronic diseases appear amongst the most frequent nowadays. Chronic conditions are defined by the World Health Organization as long-term and generally slow progressing diseases [1] that cause a major economic impact, both for health systems, as well as at family or personal levels. It is important to analyze these conditions for several reasons [2]:

(1) they account for 60% of global deaths; and (2) they trigger 75% of public health expenditure (80% of primary care consultations, and 60% of hospital admissions, are related to chronic diseases).

Essential hypertension (EH), diabetes mellitus (DM), and cardiovascular pathologies are among the most common chronicities, especially in people aged over 64 [3]. The treatment of EH and DM has become an important challenge. It is expected that the number of adults with EH will reach the figure of 1.56 billion patients [4] in 2025, and 552 million people will suffer from DM by 2030 [5]. Additionally, about 70% of diabetic patients also suffer from EH. In particular, the risk of developing diabetes is 2.4 times higher in hypertensive patients than in normotensives [6]. In this scenario, drugs also take an important role. Several studies support that the consumption of antihypertensive drugs may increase the probability of suffering from DM [7]. In addition, approximately 15–20% of hypertensive people present lower glucose tolerance than healthy patients.

The economic burden of chronic diseases demonstrates the need to address them differently, so that costs can be reduced while improving patient care and quality of life. There are several approaches in the literature for identifying patients who are at risk of suffering from certain chronic diseases. Some of these studies are based on questionnaires and do not take into account data recorded by the electronic health records (EHR) [8]. Other studies are based on information recorded from clinical examinations, which include the patient's medical history, comprising diagnoses, procedures, or drugs [9, 10]. However, these studies are limited to only one chronic condition in the population under study.

Patients in this work can have more than one chronic condition, and they are stratified by applying an internationally validated Patient Classification System (PCS) [11]. In essence, a PCS is a medical decision tree consisting of clinical rules. When PCSs consider all encounters of the patient with the health system during a period of time (usually one year), they are called Population Classification Systems, and they can be used to determine the population morbidity across time. In this work, we use the population grouper named Clinical Risk Groups (CRGs) [12, 13], since it has been internationally validated, and offers clinical relevance of the groups of patients of interest. To determine a patient's health status, the CRG use the age and gender of the patient, as well as coded information about diagnoses, procedures, and drugs. This way, each patient is assigned to a single and mutually exclusive CRG category or health status, where categories are coded as described in Sect. 2. In [14], we used the stratification provided by CRG to analyze the characteristic profile and the evolution of diagnosis codes according to the International Classification of Diseases 9th revision, Clinical Modification (ICD9-CM) [15], focusing on healthy and hypertensive patients. In addition, we proposed a classification model based on machine learning techniques (in particular, Support Vector Machines) to determine the health status of a patient based on age, gender and diagnosis codes.

Since most machine learning techniques are considered as black boxes, in this work we take a step further and explore the application of decision trees to construct classification models that clinicians can interpret. Regarding our previous study, we now increase both the number of features and chronic conditions.

Specifically, we will consider three kinds of features (demographic, diagnoses, and drugs), and at least two chronic conditions (hypertension and diabetes).

The rest of the paper is organized as follows. Section 2 describes the data analyzed in this work. Methods are explained in Sect. 3, while Sect. 4 describes the results. Finally, Sect. 5 presents a discussion and the main conclusions.

2 Data Set Description

In this work we study data from healthy and chronic patients according to the stratification provided by the CRG. Data were compiled by the University Hospital of Fuenlabrada (UHF) during the year 2012. The UHF is a public hospital of the south area of Madrid (Spain) covering residents from several cities (more than 213.000 people in total).

Each encounter of the patient with the public health system may include the following information about the patient: the date when the encounter took place, age, gender, diagnoses, procedures and pharmacological dispensing. Specifically, we worked with diagnoses collected by the hospital and associated primary care centers (coded in ICD9-CM). To reduce the number of codes, we considered just the first 3 digits, with the exception of E and M codes, which are codes of 4 characters, providing 1517 different diagnoses. We also considered pharmacological dispensing, provided by the Madrid area health institution according to the Anatomical Therapeutic Chemical (ATC) Classification System [16]. Though there are 3430 different ATC codes (coded using a combination of seven letters/digits), we have only worked with 746 by omitting the last two digits of the code, which refer to the chemical substance.

Based on those data, patients were assigned to one health status according to the CRG system, as explained in Sect. 1. There are 1080 categories of CRGs, identified by a five-digit number. The first digit refers to the core group, yielding 9 health statuses: (1) healthy, (2) history of significant acute disease, (3) single minor chronic disease, (4) minor chronic diseases in multiple organ systems, (5) significant chronic disease, (6) significant chronic diseases in multiple organ systems, (7) dominant chronic disease in three or more organ systems, (8) dominant/metastatic malignancy, and (9) catastrophic. The second, third, and fourth digits represent the base CRG and provide a total of 272 categories. The fifth digit refers to the severity level, providing a total of 1080 categories. In this paper we discarded the severity level digit in order to have a reasonable number of patients in each category of interest. Specifically, we focused on five of the 242 categories: CRG-1000, containing 46835 individuals categorized as healthy; CRG-5192, which includes 12447 patients classified as hypertensive patients; CRG-5424, with 2166 diabetic patients; CRG-6144, which contains 3179 individuals with EH and DM; and CRG-7071, with 547 patients who suffer EH, DM and other chronic illness different from cerebrovascular disease.

For each base CRG we study demographic data such as gender and age. Starting with gender, we obtained that in CRG-1000 and CRG-5192 the proportion of men and women is similar (0.49 vs. 0.51 and 0.48 vs. 0.52, respectively).

There is a higher proportion of men in CRG-5424 and CRG-6144 (0.66 vs. 0.34 and 0.55 vs. 0.45, respectively). However, the presence of women in CRG-7071 is greater (0.55 vs. 0.45), possibly because of their longer life expectancy, since patients in this category are older than patients in the rest of the considered categories. Figure 1 shows the age distribution for each category. Note that: (1) the mode of the distribution is about 60 years for categories with chronic conditions; and (2) the number of comorbidities increases with age. Furthermore, we present a descriptive analysis of both diagnosis and drug codes for each category, focusing on just those related to DM and EH.

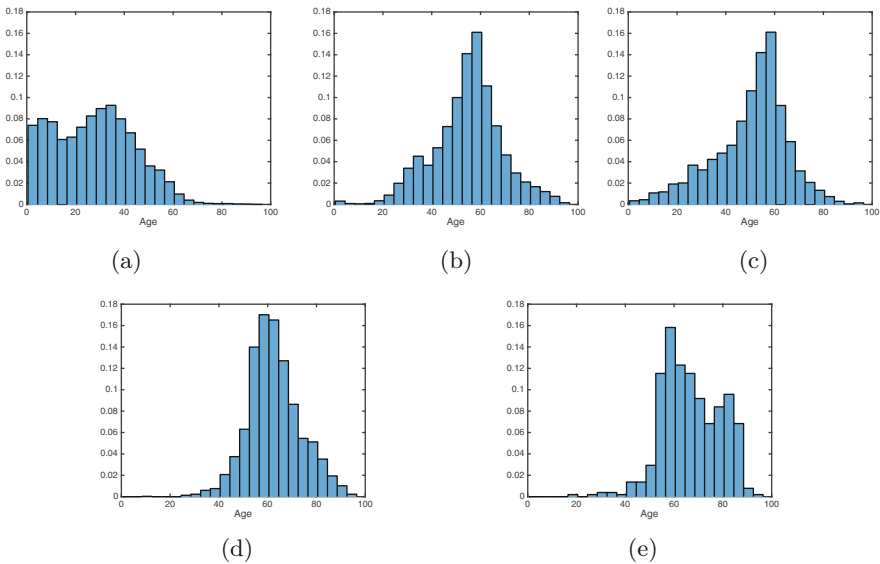


Fig. 1. Age distribution in: (a) CRG-1000, (b) CRG-5192, (c) CRG-5424, (d) CRG-6144, and (e) CRG-7071.

CRG-1000 just has healthy patients, i.e., those with no chronic health problems. However, 0.54% (0.02%) of patients have at least one diagnosis (drug) code related to DM. Likewise, 1.28% (1.59%) of healthy patients have at least one diagnosis (drug) related to EH.

CRG-5192 just contains patients classified as hypertensive by the CRG system, without another dominant chronic disease. However, about 21.5% of patients in this category have no diagnosis code associated with hypertension, and about 7.5% of patients have at least one diabetes diagnosis. In addition, 11.8% of patients do not have ATC codes related to hypertension, meaning that they are not following any treatment. This diversity in diagnoses and drugs reflects the complexity of the problem.

CRG-5424 includes patients classified as diabetics, with no other dominant chronic diseases. It is interesting to note that about 7.1% of patients in this

category have no diagnosis code related to DM, and 7.9% have no ATC codes related to diabetes treatments. Moreover, about 9.1% of patients have diagnostics related to EH, and nearly 54% take antihypertensive medication. As previously remarked, this casuistry shows the difficulty in the data.

CRG-6144 includes 3179 people classified as diabetic and hypertensive patients. Over 73% of patients have a diagnostic related to EH, and over 90% have a diagnostic related to DM. The rest of patients do not have any diagnostic related to EH and DM, although they are classified as such. Regarding drugs, more than 90% of all patients are under treatment.

CRG-7071 includes patients with DM, EH, and any dominant chronicity other than cerebrovascular disease. In this case, about 28% of all patients are classified as hypertensive/diabetic ones, but they have no diagnoses codes related to such chronic conditions. Since a clinical viewpoint, they are classified in this category because they are taking antihypertensive and/or antidiabetes drugs.

3 Methods

Machine learning techniques are used to build data-driven models. In general, they are nonparametric models and therefore do not require assumptions regarding the underlying process that generates the data. Supervised approaches are characterized because the model is learned from a set of N instances, each one with a collection of m features or attributes and the desired output. In this work, instances correspond to patients. Each patient is characterized by different kinds of attributes: gender (1), age (1), diagnosis codes (1517), and drug codes (746). The desired output is the health status or category provided by the CRG when using all encounters with the health system in 2012.

The goal is to construct a classifier able to represent interactions among attributes and classes while providing good generalization [17]. Since there is a plethora of machine learning techniques [17], we focus in this work on tree-based methods (decision trees) because they are easy enough to understand by human experts and have been successfully used in various fields of application.

A decision tree provides a tree-like graph model for performing classification tasks, where each node encodes a simple rule that is easy to interpret by end users. By visualizing a decision tree, clinicians can obtain new medical insights related to the discriminative power of the features and their relationships.

According to the machine learning terminology, any tree is composed of three kind of nodes: root, internal, and leaf nodes. The root node is the origin of the tree, the internal nodes correspond to conditions applied on the attributes, and the leaf nodes represent the classes. The branches are segments joining the nodes. The number of internal and leaf nodes is related to the tree size (complexity), and with the accuracy and interpretability of the model. The higher the number of internal nodes, the more difficult it will be to analyze and interpret the tree. According to the parsimonious principle (Occam's razor, see [18]) simpler models are preferred over more complex ones. The tree complexity can be controlled by using pruning techniques.

In this work we have used the C4.5 algorithm [19], one of the first and simplest algorithms for constructing decision trees. It is based on the ‘divide and conquer’ method: if all of the instances in a node belong to the same class, or there are just a few instances, the node is a leaf node and is labeled according to the majority class; otherwise, the node is converted into an internal node, and two branches are created recursively. The final classifier consists of a collection of rules defined by the path from the root to a leaf node. Each rule is encoded in an internal node and considers just one attribute. The particular attributes appearing at each internal node depends on the concept of information gain (entropy). The idea is to split branches by choosing the attribute that provides the highest information gain ratio [19]. To avoid overfitting, a pruning stage following a bottom-up approach is applied to get the final model.

To control the tree expansion, there are some hyperparameters to take into account. The first one is the minimum number of instances in the two most popular branches of a node (M): the noisier the data, the higher M should be. The second hyperparameter is the confidence value C , which is related to the running stage: lower values for C incur in heavier pruning.

As in many real-life problems (especially in medicine), class imbalance can affect learning. As suggested by [20], we followed an undersampling strategy in order to balance the classes, by removing instances from the majority classes.

4 Results

In this section we describe the experimental setup and results provided by decision trees when applied to the data presented in Sect. 2.

4.1 Experimental Setup

We considered two sets of features: (FS-1) demographic data (age and gender) and ICD9 codes (diagnoses); and (FS-2) demographic data, ICD9 codes, and ATC codes (drugs dispensation). For each set of features, we constructed a decision tree to discriminate patients in categories related to hypertension (CRGs 5192-6144-7071), and another tree that also includes healthy patients (CRG-1000). We performed a similar analysis for examining the evolution of diabetic patients (CRGs 5424-6144-7071, and CRGs 1000-5424-6144-7071). Since the categories were imbalanced (see Sect. 2) we employed an undersampling strategy and created a balanced dataset, whose size is limited by the class with less instances (CRG-7071). We carried out 50 trials in which we randomly chose 80% of the data for training, and the rest for testing.

Tuning free parameters. As described in Sect. 3, the free parameters are the minimum number of instances M , and the confidence value C . Classification results were obtained in terms of accuracy, kappa value [21], tree size, and number of leaves in the tree. We explored several values: $M = \{1, 5, 10, 15, 20, 25, 30, 40\}$, and $C = \{0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5\}$. We selected the pair M - C that maximizes both the accuracy and kappa value, and minimizes the size of the tree.

4.2 Decision Trees

We explored the potential to predict the evolution of diabetes and hypertension when considering the two sets of features described in Sect. 4.1. Firstly, we analyzed the performance in terms of accuracy and kappa values (mean and standard deviation, see Table 1). Subsequently, we drew clinical conclusions regarding the features considered in each decision tree. Starting with the set of categories related to hypertension (CRGs 5192-6144-7071), the average values for accuracy and kappa were 85.40% and 0.78, respectively, when considering FS-2. These values were lower when drugs were not included, i.e., when using FS-1. Note that when discriminating patients with EH-DM and just one of these chronic conditions (EH or DM), a better performance is obtained for EH (CRGs 5192-6144-7071) versus DM (GRGs 5424-6144-7071). This may be because most of patients classified as diabetic also take antihypertensive drugs, which can be misleading since these drugs can be used to treat several diseases.

When a patient is seen for consultation (especially in primary care), it could happen that the diagnosis code that is registered is the one corresponding to the purpose of the visit, while other pathologies would be left without coding. In this scenario, the ATC codes may be very useful for identifying chronic patients. From now on, the results provided are according to FS-2.

As expected, better results were obtained when healthy patients were considered as a fourth category to discriminate (see Table 1). The worst results were obtained when discriminating categories related to diabetic patients. Thus, the average accuracy was 82.55% (73.31%), and the kappa value was 0.77 (0.64) considering FS-2 (FS-1) and healthy patients. When healthy patients were not considered, performance decreased to 80.14% (68.24%) for accuracy, and to 0.70 (0.52) for kappa, with FS-2 (FS-1).

Table 1. Mean and standard deviation for classification accuracy (first row in the cell) and kappa value (second row in the cell), when discriminating among different categories (columns), using two sets of features (rows).

Features	CRGs 1000, 5192, 6144, 7071	CRGs 5192, 6144, 7071	CRGs 1000, 5424, 6144, 7071	CRGs 5424, 6144, 7071
Gender, age and ICD9 codes	79.04 (1.83)	75.42 (1.95)	73.31 (2.03)	68.24 (2.21)
	0.72 (0.02)	0.63 (0.03)	0.64 (0.03)	0.52 (0.03)
Gender, age, ICD9 and ATC codes	86.61 (1.58)	85.40 (2.01)	82.55 (1.93)	80.14 (1.89)
	0.82 (0.02)	0.78 (0.03)	0.77 (0.03)	0.70 (0.03)

Figure 2 shows the trees obtained when only considering CRGs 1000-5192-6144-7071 (a) and CRGs 1000-5424-6144-7071 (b). Ovals refer to input features (the codes have tree digits when referring to diagnoses, and five when referring

to drugs), and rectangles to categories. We analyzed in detail the set of features selected by each decision tree for discriminating patients into different categories. Table 2 shows the features (rows) selected in the root and internal nodes when considering classification into different classes (columns). It is important to note that the presence of diagnoses codes ‘250’ (DM) and ‘401’ (EH) is consistent with the categories to be discriminated. Thus, both diagnosis codes appear when discriminating healthy patients versus chronic patients who suffer from one or both conditions (hypertension/diabetes). In contrast, the code ‘250’ does not appear when only discriminating chronic patients of CRGs 5424-6144-7071. The reason is that patients in these three categories are diabetic, and the code ‘250’ would not be relevant for discriminating among these groups. A similar comment applies to the presence of code ‘401’ when only CRGs 5192-6144-7071 are considered. This result provides clinical significance and offers us a tool to study the evolution of the chronic conditions from a statistical and clinical perspective.

Regarding the diagnosis code ‘496’ (chronic obstructive pulmonary disease), it is likely to appear in the model for discriminating patients in CRG-7071. It is interesting to remark that the code ‘571’ (chronic liver disease and cirrhosis) is discriminative in all trees except when all of the classes in the tree consider diabetes. The association between cirrhosis and diabetes has been reported by other studies [22], which support that mortality from cirrhosis is nearly three times higher in patients with DM than in the general population. So, it is likely that the ‘571’ diagnosis code is so frequent in diabetic patients that it does not provide useful information to differentiate among the corresponding categories.

Attributes related to drugs appearing in the tree are also important from a clinical viewpoint. Thus, even when the antihypertensive drugs (ATC codes ‘C03CA’, ‘C07AB’, ‘C08CA’, ‘C09AA’ and ‘C09CA’) can be indicated for other pathologies different from DM and EH (e.g. cardiac pathologies), some of these drugs appear in all trees except when discriminating categories from GRGs 5192-6144-7071. The reason is that all patients are hypertensive, so the consumption of antihypertensive drugs cannot help to discriminate the categories.

The ATC code ‘A10BA’ (blood glucose lowering drugs, excluding insulins) is discriminative when the diagnosis code ‘250’ appears in the tree, but not in the patient registry (see Figs. 2 and 3). This means that when the diagnosis code ‘250’ was not recorded, the code ‘A10BA’ provides information to discriminate between the categories. According to clinicians, the reason is that this drug is only prescribed for diabetics (it is not used to treat other diseases).

Another therapeutic group (ATC code ‘B01AB’), which comprises antithrombotic agents and heparines, has indications in situations of high risk of blood clots, leading to severe pathologies related to the cardiovascular and cerebral system. This drug is prescribed to patients with significant chronic conditions, being the alteration of the cardiac rhythm (atrial fibrillation) the most frequent one. According to the tree, the presence of this drug helps to discriminate between CRG-6144 and CRG-7071.

The ATC code ‘N05AL’ (antipsychotics) appears in all trees and allows to classify the patient into CRG-7071. It is also remarkable that this code appears

Table 2. List of features selected by the tree (x) when discriminating among different categories (the four columns on the right).

Feature	Description	1000, 5192, 6144, 7071	5192, 6144, 7071	1000, 5424, 6144, 7071	5424, 6144, 7071
250	Diabetes mellitus	x	x	x	
401	Essential hypertension	x		x	x
496	Chronic airway obstruction, not elsewhere classified	x	x	x	x
571	Chronic liver disease and cirrhosis	x	x	x	
789	Other symptoms involving abdomen and pelvis				x
784	Symptoms involving head and neck				x
A10AE	Insulins and analogues for injection, long-acting			x	x
A10BA	Blood glucose lowering drugs, excl. Insulins	x	x	x	
A10BD	Combinations of oral blood glucose lowering drugs		x		
B01AB	Heparin group. Antithrombotic agents	x	x	x	x
C03CA	Sulfonamides, plain. High-ceiling diuretics			x	x
C07AB	Beta blocking agents, selective	x		x	x
C08CA	Selective calcium channel blockers with mainly vascular effects. Dihydropyridine derivatives			x	x
C09AA	ACE inhibitors, plain	x		x	x
C09CA	Angiotensin II antagonists, plain				x
C10AA	HMG coa reductase inhibitors			x	
H02AB	Glucocorticoids	x	x		x
M01AE	Antiinflammatory and antirheumatic products, non-steroids. Propionic acid derivatives	x			
N02BB	Other analgesics and antipyretics. Pyrazolones				x
N02BE	Other analgesics and antipyretics. Anilides		x		
N04BC	Dopaminergic agents. Dopamine agonists	x	x	x	x
N05AL	Antipsychotics. Benzamides	x	x	x	x
N05AH	Antipsychotics. Diazepines, oxazepines, thiazepines and oxepines	x			x
N06AX	Antidepressants	x	x	x	x
N06DA	Antidepressants	x	x	x	x
R03AC	Drugs for obstructive airway diseases. Selective beta-2-adrenoreceptor agonist				x
Age		x	x	x	x

in the root node (of the tree associated to CRGs 5424-6144-7071, see Fig. 3(b)), or just after it when the diagnosis code ‘250’ is among the diagnoses of the patient (rest of trees). This result is in line with the evidenced relationship between diabetes and depression [23]. For the rest of antipsychotic and antidepressant drugs (starting with ‘N’), their consumption always helps to discriminate patients in CRG-7071. It is worth recalling that CRG-7071 encompasses more chronic conditions, which generate more psychiatry and behavioral manifestations (e.g. changes in the sleep habits) that may determine the use of these drugs.

It is also interesting to point out that the ATC code ‘R03AC’ allows to discriminate patients in CRG-7071 when the diagnosis code ‘496’ does not appear in the classification tree for CRGs 5424-6144-7071 (see Fig. 3(b)). This is another example where drugs can be informative when the diagnosis code directly linked to the drug is not present.

Regarding age, it appears in all the trees and always very close to the leaf nodes. From Figs. 2 and 3 it seems that age is discriminative when the diagnosis code ‘250’ is not present: young patients correspond to CRG-1000, while older patients correspond to the category with more chronic conditions (CRG-7071).

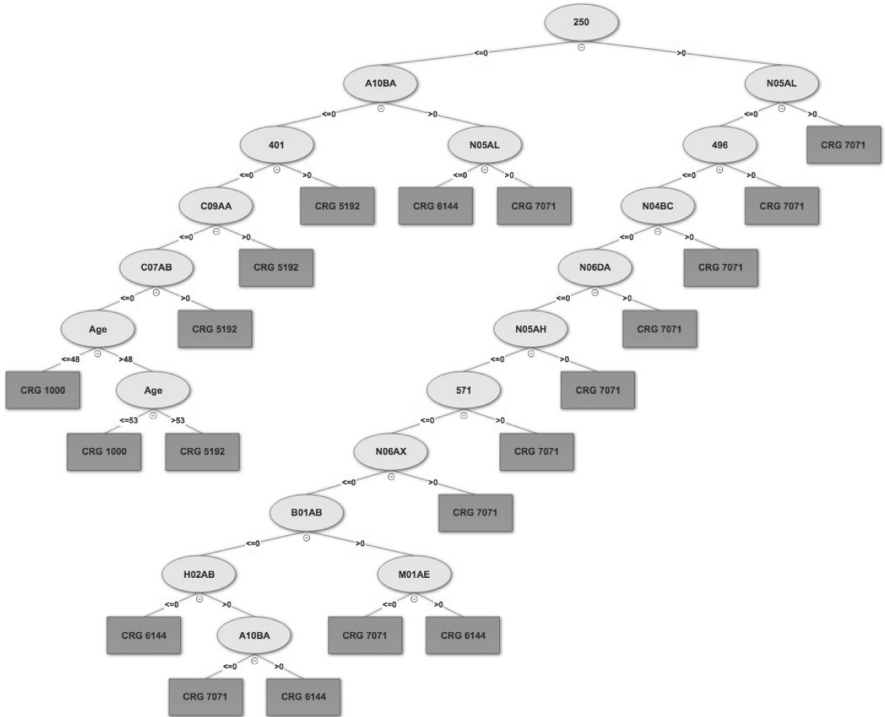
For the tree associated with CRGs 5424-6144-7071 (see Fig. 3(b)), where code ‘250’ (DM) is not discriminative, the feature age appears in the two branches generated from the diagnosis code ‘401’ (EH). Thus, if the code ‘401’ is present, the age allows us to discriminate between CRG-6144 and CRG-7071 (oldest patients). However, if the code ‘401’ is not present, age allows to discriminate between CRG-5424 and CRG-7071. These results are supported by clinical knowledge.

5 Discussion and Conclusions

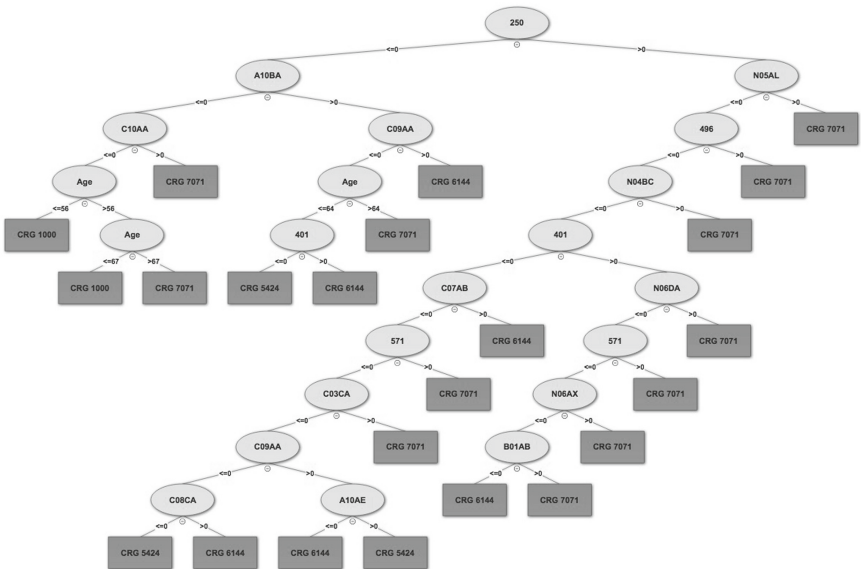
As a consequence of aging, the prevalence of chronic pathologies increases. In the analyzed dataset in this work, EH and DM are the most frequent chronic conditions. Specifically, in Spain hypertension affects more than 50% of the population over 64 years of age, and diabetes to almost 20%. In addition, these diseases represent a great economic expense, both to the health system and to the individual. This study verifies the relationship between the number of dominant chronic conditions and the age of the patients.

Decision trees are designed to classify patients into different categories of chronic conditions according to the PCS called CRG. As might be expected, the performance is always better when healthy patients are considered, since they are easier to be distinguished from the rest of patients suffering from chronic conditions. Regarding predictive analyses, when discriminating between patients with EH-DM and just one of these chronic conditions, a better accuracy is obtained for EH (85.4%) versus DM (80.1%).

The resulting trees provided interpretable models in accordance with clinical knowledge. In fact, the trees showed a relationship between EH/DM and

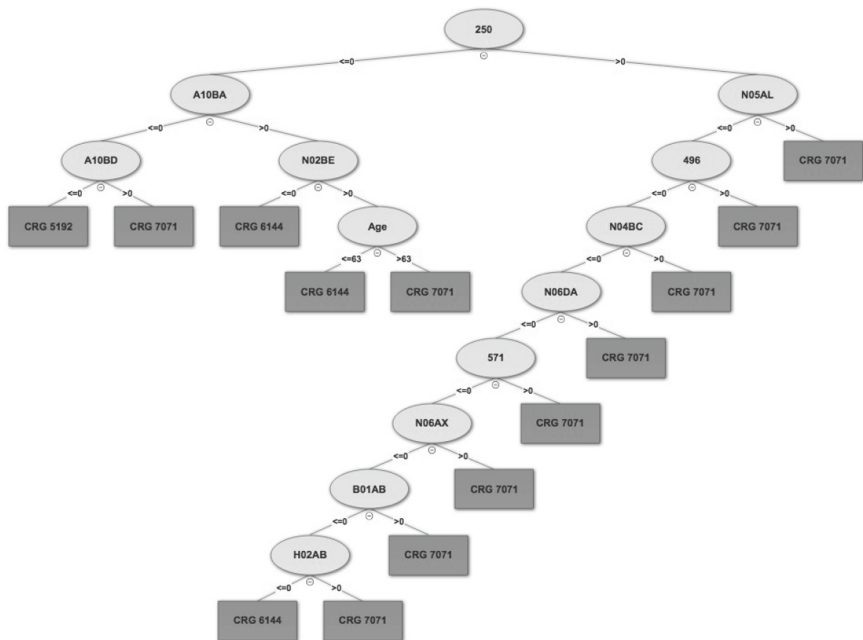


(a)

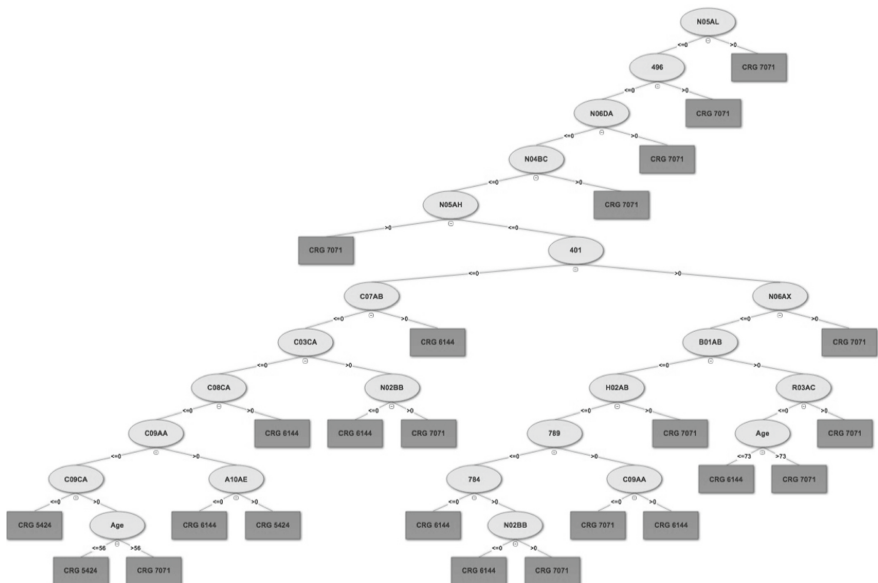


(b)

Fig. 2. Classification trees considering age, gender, diagnoses codes (nodes with three digits) and drug codes: (a) CRGs 1000-5192-6144-7071; (b) CRGs 1000-5424-6144-7071.



(a)



(b)

Fig. 3. Classification trees considering age, gender, diagnoses codes (nodes with three digits) and drug codes: (a) CRGs 5192-6144-7071; (b) CRGs 5424-6144-7071.

diseases/drugs related to other chronic conditions, such as those of the respiratory system and mental disorders. Lastly, drugs can be informative when the diagnosis code related to the chronic condition is not registered.

We can conclude that a single year is a short time to extract distinctive patterns of the patients' health status. As a future work, we will analyze the temporal evolution of diagnoses and drugs considering chronic patients across several years, complementing this study with both structured and unstructured data.

Acknowledgments. This work has been partly supported by Research Projects TEC2016-75361-R, TIN2014-62143-EXP, and TIN2015-70799-R from the Spanish Government, and Research Project DTS17/00158 from Carlos III Institute.

References

1. WHO: OMS (2017). http://www.who.int/topics/chronic_diseases/es/
2. Anderson, G., Horvath, J.: The growing burden of chronic disease in America. *Public Health Rep.* **119**(3), 263–270 (2004)
3. Soni, A., Mitchell, E.: Expenditures for commonly treated conditions among adults age 18 and older in the US civilian noninstitutionalized population, 2013 (2001)
4. Kearney, P.M., et al.: Global burden of hypertension: analysis of worldwide data. *Lancet* **365**(9455), 217–223 (2005)
5. International Diabetes Federation: IDF Diabetes Atlas, 5th edn. International Diabetes Federation, Brussels (2011)
6. Hollis, W.C.: Aggravation of diabetes mellitus during treatment with chlorothiazide. *JAMA* **176**(11), 947–949 (1961)
7. Karnes, J.H., Cooper-DeHoff, R.M.: Antihypertensive medications: benefits of blood pressure lowering and hazards of metabolic effects. *Expert Rev. Cardiovasc. Ther.* **7**(6), 689–702 (2009)
8. Smith, S., Holohan, J., McAuliffe, A., Firth, R.: Irish diabetes detection programme in general practice. *Diabet. Med.* **20**(9), 717–722 (2003)
9. Greaves, C., et al.: A simple pragmatic system for detecting new cases of type 2 diabetes and impaired fasting glycaemia in primary care. *Fam. Pract.* **21**(1), 57–62 (2004)
10. Woolthuis, K., et al.: Identifying people at risk for undiagnosed type 2 diabetes using the GP's electronic medical record. *Fam. Pract.* **24**(3), 230–236 (2007)
11. Davis, N.A., LaCour, M.: *Foundations of Health Information Management*. Elsevier Health Sciences, Amsterdam (2016)
12. Averill, R.F., et al.: Development and evaluation of clinical risk groups (CRGs). *3M Health Information Systems* (1999)
13. Hughes, J.S., et al.: Clinical risk groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management. *Med. Care* **42**(1), 81–90 (2004)
14. Fernández-Sánchez, J., et al.: Clinical risk groups analysis for chronic hypertensive patients in terms of ICD9-cm diagnosis codes. In: *Proceedings of the 4th International Conference on Physiological Computing Systems - Vol. 1. INSTICC*, SciTePress, pp. 13–22 (2017)
15. Centers for Disease Control and Prevention: *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* (2011)

16. WHO: collaborating centre for drug statistics methodology. In: Guidelines for ATC Classification and DDD Assignment (2016)
17. Duda, R.O., et al.: Pattern Classification, 2nd edn. Wiley, New York (2001)
18. Ripley, B.D.: Pattern Recognition and Neural Networks. CUP, Cambridge (1996)
19. Quinlan, J.: C4.5. Programs for Machine Learning. Morgan Kaufmann, Burlington (2014)
20. Wallace, B.C., et al.: Class imbalance, redux. In: Proceedings of the International Conference on Data Mining, Vancouver, Canada, pp. 754–763, 11–14 December 2011
21. Ferri, C., et al.: An experimental comparison of performance measures for classification. Pattern Recogn. Lett. **30**(1), 27–38 (2009)
22. Simó, R., et al.: High prevalence of hepatitis c virus infection in diabetic patients. Diabetes care **19**(9), 998–1000 (1996)
23. Colunga-Rodríguez, C., et al.: Type 2 diabetes and depression in Guadalajara, Mexico, 2005. Revista de salud publica **10**(1), 137–149 (2008)

Biomedical Image Analysis



Stochastic Geometry for Automatic Assessment of Ki-67 Index in Breast Cancer Preparations

Marek Kowal¹(✉), Marcin Skobel¹, Józef Korbicz¹, and Roman Monczak²

¹ Faculty of Computer, Electrical and Control Engineering,
Institute of Control and Computation Engineering,
University of Zielona Góra, ul. Licealna 9, 65-417 Zielona Góra, Poland
m.kowal@issi.uz.zgora.pl

² Department of Pathomorphology, University Hospital in Zielona Góra,
ul. Zyty 26, 65-046 Zielona Góra, Poland

Abstract. Proliferative activity of cells is one of the most critical factors in breast cancer diagnosis. It is used to evaluate tumor cell progression and to predict treatment responses in chemotherapy. Ki-67 is a nuclear biomarker commonly used to measure cellular proliferation rate. The ratio between the number of Ki-67 positive tumor nuclei and all tumor nuclei defines Ki-67 index. However, manual cell counting is tedious and time consuming because hundreds of nuclei must be labeled. To speed up the analysis process, nuclei can be segmented automatically and then classified based on staining color. Unfortunately, segmentation of individual nuclei is a big challenge because they often create complex clusters comprised of many touching and overlapping nuclei. To deal with complexities and ambiguities of cytological material we propose a generative model which approximates nuclei using ellipses. We assume that the process of generating a cytological sample has stochastic nature. Therefore it is possible to reconstruct this process using marked point process tuned according to observed cytological sample. To verify the potential of the proposed method, we applied it to determine Ki-67 index in breast cancer immunochemistry samples. The results of experiments have shown that Ki-67 indices determined by proposed approach correlate well with those computed manually.

Keywords: Breast cancer · Ki-67 index · Nuclei segmentation
Stochastic process · Steepest ascent optimization
Maximum a posteriori estimation

1 Introduction

Biological material necessary for immunochemical examination is taken from affected breast tissue using needle biopsy. The material is then subjected to an immunochemical processing, fixing and staining. Ki-67 positive nuclei are stained

by diaminobenzidine (DAB) and have a brown color; negative nuclei appear as blue due to hematoxylin (H) dye. Finally, slide glass with stained material is scanned and evaluated by the pathologist. Ki-67 index is calculated as a ratio between brown-colored nuclei and all nuclei. Individual cases of breast cancer are classified using a fixed cut-off point for the Ki-67 index. During the 2011 St. Gallen International Breast Cancer Conference agreed a cut-off value of 14% based on the study by Cheang *et al.* [5]. The recommendation had changed in 2013 when most of the panels of St.Gallen conference suggested a threshold of 20%. Finally, the 2015 consensus suggested that cut-off value should be tuned separately for each laboratory based on median value from local results.

The main problem when determining the Ki-67 index is that many pathologists do not have time to visually count the cell nuclei. This process is extremely time-consuming because it requires choosing the most active areas in proliferative rate (hotspots) and then pathologists have to mark manually few hundreds of nuclei. Counting by visual inspection is highly subjective because it is linked to interpretation. Many studies report that inter-observer variability in determining the Ki-67 index is high [16].

Automatic nuclei counting would have allowed avoiding tedious and in fact unnecessary work for pathologist. Also, the process would become free from the subjective assessment of the pathologist. However, we must remember that the accuracy of nuclei segmentation is critical to performance of Computer-Aided Cytology (CAC). Unfortunately, immunochemical images of cytological material are rather challenging for existing object detection methods. In hotspots, they usually create complex, random and heterogeneous structures like clumps and nests. Nuclei are often overlapping and touching. A number of scientific centers conduct an intensive research to develop efficient algorithms for object segmentation. The most common approaches are based on the image thresholding, color deconvolution, data clustering, watershed, active contours and deep learning [4, 8]. One of the well-known systems for determining Ki-67 index is ImmunoRatio [15]. This system estimates the Ki-67 index as the ratio of the brown-colored area of the image to the brown- and blue-colored area of the image. Staining components (DAB and H) are separated using H+DAB color deconvolution technique and intensity thresholding [12]. For the ImmunoRatio to be effective, it must be appropriately calibrated for data from a specific laboratory. There are a lot of reports in the literature about the high accuracy of this system [18]. Saha *et al.* proposed deep learning framework for hotspots detection and proliferation scoring [13]. They achieved 93% precision, 0.88% recall and 0.91% F-score value. In addition, in their work, we can find a comprehensive list of methods used for estimating the Ki-67 index. In [1] a classifier based algorithm was proposed for Ki-67 scoring in BC tissue microarray images. Proposed approach reached 90% classification accuracy with 0.64 kappa concordance. Xing *et al.* applied boundary delineation algorithm to localize tumor and nontumor cells for automatic Ki-67 counting [17]. Their algorithm shows promising performance measure in comparison with other popular Ki-67 scoring techniques. Here, we can also find an overview of the other methods used to estimate Ki-67 index.

The aim of our research is to develop and test the new method of Ki-67 scoring based on stochastic geometry. Stochastic geometry is a branch of probability theory that deals with the analysis of random spatial patterns [6]. They are successfully applied wherever there are heterogeneous structures whose spatial distribution is random [7, 10]. We can observe such distribution for cytological material in immunochemical images.

The idea of our approach is to transform H+DAB stained image into a model build up with ellipses which vary in the location, size and orientation. The goal is to find such configuration of ellipses that fit the image data best without violating a priori preferences regarding the distribution of nuclei. The next step is to divide ellipses into two classes due to the color of actual nuclei. This step was carried out with the help of color deconvolution, which allowed us to discover DAB staining areas and H staining areas. The ellipses are by default classified to blue-colored (H) objects and then they are changing their assignment if they are touching the brown-colored (DAB) area. Finally, nuclei models are counted in order to determine Ki-67 index.

To verify the effectiveness of proposed approach, it was applied to estimate Ki-67 index for cytological images of breast cancer. The obtained results were compared with the reference results obtained for manual segmentation and with the results coming from ImmunoRatio.

The remainder of this paper is organized as follows. Section 2 gives the description of methods applied to determine Ki-67 index. Results of experiments are presented in Sect. 3. Concluding remarks are given in Sect. 4.

2 Method

2.1 Method Overview

Before it is possible to build immunochemical models of actual images, pre-processing must be applied to these images to separate the areas stained by H and DAB. To achieve this, the RGB image is subjected to H+DAB deconvolution using the procedure implemented in ImageJ plugin [12]. Three separate intensity images are created as a result. The first represents how much of H has been deposited in nuclei, second how much of DAB has been deposited in nuclei, and third residuals (see Fig. 1). For further processing, we are using the image of H density and DAB density. Next step is carried out to determine binary masks of H area and DAB area. Both masks are determined based on intensity thresholding. The H mask is combined with the DAB mask using AND logic operator. As an effect, we are obtaining H+DAB mask used by stochastic based algorithm to detect elliptical objects which resemble the nuclei. The DAB mask alone cannot be used to segment Ki-67 positive nuclei because in some cases only small part of the nucleus is colored by DAB but anyway this means that the nucleus should be counted as Ki-67 positive. Therefore, H+DAB mask is used to segment all nuclei using stochastic geometry approach, and then DAB mask is used to indicate Ki-67 positive nuclei by checking which ellipses overlaps with the area of the DAB mask. The entire procedure is illustrated in Fig. 1.

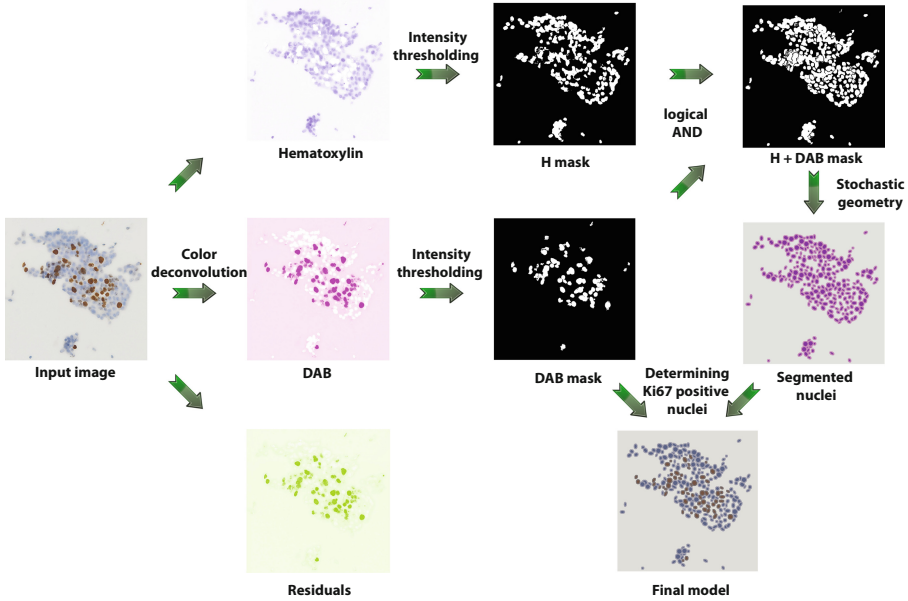


Fig. 1. Overview of the method

The crucial step of the proposed strategy is to find a proper configuration of ellipses that fit the H+DAB mask. Stochastic geometry based approach is trying to cover as many as possible pixels of nuclei area on H+DAB mask using a collection of ellipses of different size and orientation [2, 9]. At the same time, the algorithm is trying to avoid to cover background pixels. Of course, there are many different collections of ellipses that fit well to the given H+DAB mask. To choose the best configuration, we need a prior knowledge about preferable configurations of nuclei. For this purpose, we control pairwise interactions between ellipses to limit the number of overlaps between objects. Thus, configurations with a large number of overlapping ellipses are less likely than those that have less overlaps.

2.2 Marked Point Process

In the considered approach nuclei segmentation in H+DAB image boils down to finding the configuration of ellipses varying in size, location and orientation which covers precisely the H+DAB mask and follows the distribution given by aprior model. To find such a configuration of ellipses, we had to assume that the image generation process can be described using marked point process (MPP). So the crucial element necessary to reconstruct the input image is the knowledge of conditional probability mass function (pmf) $p(\mathbf{x}|\mathbf{y})$ which governs such process:

$$p(\mathbf{x}|\mathbf{y}) \propto f(\mathbf{y}|\mathbf{x})p(\mathbf{x}), \quad (1)$$

where likelihood term $f(\mathbf{y}|\mathbf{x})$ evaluates the consistency of ellipse configuration $\mathbf{x} = \{x_1, \dots, x_n\}$ respect to H+DAB mask \mathbf{y} and a prior term $p(\mathbf{x})$ reflects constraints on pairwise interactions between ellipses within configuration \mathbf{x} [2]:

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{x_i \sim x_j} h(x_i, x_j), \tag{2}$$

where $\alpha, \beta > 0$ are constants, $n(\mathbf{x})$ is the number of disks in configuration \mathbf{x} , h is the interaction function and \sim is a symmetric and reflexive relation describing ellipse overlaps.

If we assume that variables representing mask values $y_t \in \{0, 1\}$ are conditionally independent given configuration \mathbf{x} then $f(\mathbf{y}|\mathbf{x})$ takes the following form:

$$f(\mathbf{y}|\mathbf{x}) = \prod_{t \in S(x)} b(y_t; p_N) \prod_{t \in S \setminus S(x)} b(y_t; p_B), \tag{3}$$

where $b(y_t; p_N)$ and $b(y_t; p_B)$ are Bernoulli pmf's:

$$b(y_t; p_N) = \begin{cases} 1 - p_N & \text{if } y_t = 0 \\ p_N & \text{if } y_t = 1, \end{cases} \tag{4}$$

$$b(y_t; p_B) = \begin{cases} 1 - p_B & \text{if } y_t = 1 \\ p_B & \text{if } y_t = 0. \end{cases} \tag{5}$$

They are used to evaluate the likelihood of pixels on H+DAB mask \mathbf{y} within nuclei region $S(x)$:

$$S(x) = \bigcup_{i=1}^n S(x_i), \tag{6}$$

and background region $S \setminus S(x)$ respectively, where S is a pixel lattice of mask H+DAB and $S(x_i)$ is the silhouette of the ellipse x_i . Parameter p_N describe a probability of occurring actual nuclei pixel within nuclei region $S(x)$ on H+DAB mask \mathbf{y} , and p_B is a probability of occurring actual background pixel within background $S \setminus S(x)$. Both parameters were chosen arbitrarily and are shown in Fig. 2.

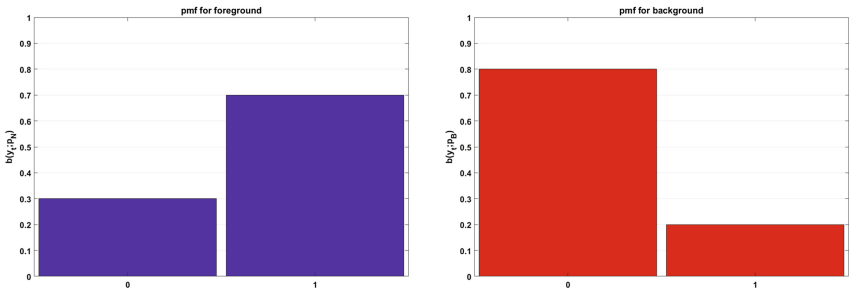


Fig. 2. Probability distributions for $S(x)$ and $S \setminus S(x)$ regions

Finally, we choose Strauss process to implement interaction model [3, 14]:

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{r(\mathbf{x})}, \quad (7)$$

where $\alpha, \beta, \gamma > 0$ are constants, $n(x)$ is the number of ellipses in configuration and $r(\mathbf{x})$ is the number of pairwise overlaps in configuration \mathbf{x} . For $0 < \gamma < 1$, model exhibits repulsive forces between ellipses and this prevent an excessive number of overlaps in ellipse configurations.

2.3 Optimization

Finding proper configuration of ellipses boils down to optimization problem where Ω is a set of all possible configurations \mathbf{x} . To solve the problem of nuclei segmentation, we must find in Ω a configuration \mathbf{x} that fit the image best without contravene a prior interaction constraints. In Bayesian framework this problem can be viewed as a maximum a posterior estimation problem:

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} f(\mathbf{y}|\mathbf{x})p(\mathbf{x}). \quad (8)$$

Unfortunately, direct sampling from $f(\mathbf{y}|\mathbf{x})$ and $p(\mathbf{x})$ is not straightforward. But, the problem becomes much more tractable if we deal with the following proportion [2, 9]:

$$\begin{aligned} w &= \ln \left(\frac{f(\mathbf{y}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1})}{f(\mathbf{y}|\mathbf{x}_k)p(\mathbf{x}_k)} \right) \\ &= \sum_{t \in S_N} \left(\ln(b(y_t; p_N)) - \ln(b(y_t; p_B)) \right) \\ &\quad + \ln(\gamma)(r(\mathbf{x}_{k+1}) - r(\mathbf{x}_k)) + \ln(\beta), \end{aligned} \quad (9)$$

where \mathbf{x}_k is the current configuration, \mathbf{x}_{k+1} is the new prospective configuration and $S_N = (S(\mathbf{x}_{k+1}) \cup S(\mathbf{x}_k)) \setminus (S(\mathbf{x}_{k+1}) \cap S(\mathbf{x}_k))$. If we limit the ways the new configurations \mathbf{x}_{k+1} can emerge by allowing only to add single ellipse u or delete single ellipse u from the current configuration \mathbf{x}_k then it becomes possible to apply steepest ascent procedure to find the local maximum. Algorithm is always choosing new configuration \mathbf{x}_{k+1} to maximize proportion w . Therefore, probability never decreases at any stage and eventual convergence is guaranteed. However, algorithm usually stuck in nearest local maxima. The pseudocode of this procedure is presented in Algorithm 1.

2.4 Ki-67 Scoring

Ki-67 index is computed as the ratio between brown-colored nuclei and all nuclei. Steepest ascent procedure segments nuclei and returns the configuration of ellipses that approximates actual nuclei but it is not able to distinguish Ki-67 positive nuclei form Ki-67 negative nuclei. To tackle this problem, we are

```

Initialization:  $\mathbf{x}_0 = \emptyset$ ;
for  $i = 0, 1, \dots$ ; do
  for All  $u$  do
    if  $u \notin \mathbf{x}_k$  then
       $\mathbf{x}_{test} = \mathbf{x}_k \cup u$ ;
       $w(u) = \ln \left( \frac{f(\mathbf{y}|\mathbf{x}_{test})p(\mathbf{x}_{test})}{f(\mathbf{y}|\mathbf{x}_k)p(\mathbf{x}_k)} \right)$ ;
    else if  $u \in \mathbf{x}_k$  then
       $\mathbf{x}_{test} = \mathbf{x}_k \setminus u$ ;
       $w(u) = \ln \left( \frac{f(\mathbf{y}|\mathbf{x}_{test})p(\mathbf{x}_{test})}{f(\mathbf{y}|\mathbf{x}_k)p(\mathbf{x}_k)} \right)$ ;
    end
    find ellipse  $u_{max}$  which maximizes  $w(u)$ ;
    if  $w(u) > 0$  then
      if  $u_{max} \notin \mathbf{x}_k$  then
         $\mathbf{x}_{k+1} = \mathbf{x}_k \cup u_{max}$ ;
      else if  $u_{max} \in \mathbf{x}_k$  then
         $\mathbf{x}_{k+1} = \mathbf{x}_k \setminus u_{max}$ ;
      else
        Return  $\mathbf{x}_k$ ;
    end
  end

```

Algorithm 1. Nuclei segmentation using Steepest Ascent procedure

marking all ellipses which are within the DAB mask as Ki-67 positive (see Fig. 1). Finally, we can quickly count all nuclei by checking how many nuclei are in the found configuration $\hat{\mathbf{x}}$. The estimate of Ki-67 index is computed as the number of positive ellipses to the number of all ellipses.

3 Results

The proposed approach was applied to estimate Ki-67 index in 20 test cases of breast cancer. Immunochemical examinations were obtained for 20 patients from the University Hospital in Zielona Góra, Poland. Immunochemical slides were digitized into virtual slides using the Olympus VS120 Virtual Microscopy System. Selected fragments (size 500×500 pixels) of these slides were used in experimental studies (see Fig. 3). Each test image contains from 100 to 400 nuclei. All images were manually marked to get the reference results of Ki-67 scoring. The accuracy of the proposed method was compared with the accuracy of ImmunoRatio system.

To compute Ki-67 index we need segmented nuclei. In our approach stochastic geometry is responsible for extracting nuclei models in the form of ellipses. The method requires an atlas of predefined models of nuclei. In this experiment, we have used atlas comprised of 240 ellipses which vary in sizes and orientations. Each ellipse is described using 3 parameters: major axis length $r_M \in [6, \dots, 15]$, minor axis length to major axis length ratio $r_R \in [0.5, 0.65, 0.8, 1]$ and orientation $\phi \in [0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ]$. The other parameters that must be defined for

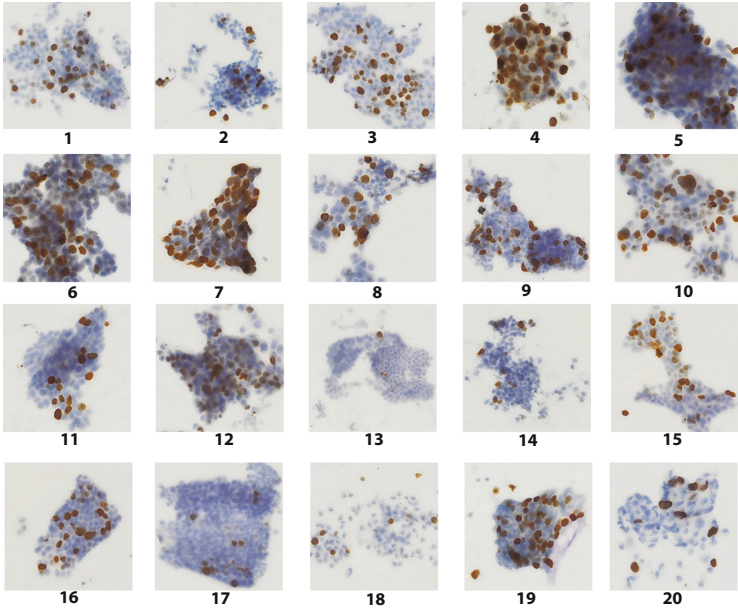


Fig. 3. Input images (Ki-67 positive nuclei have darker color)

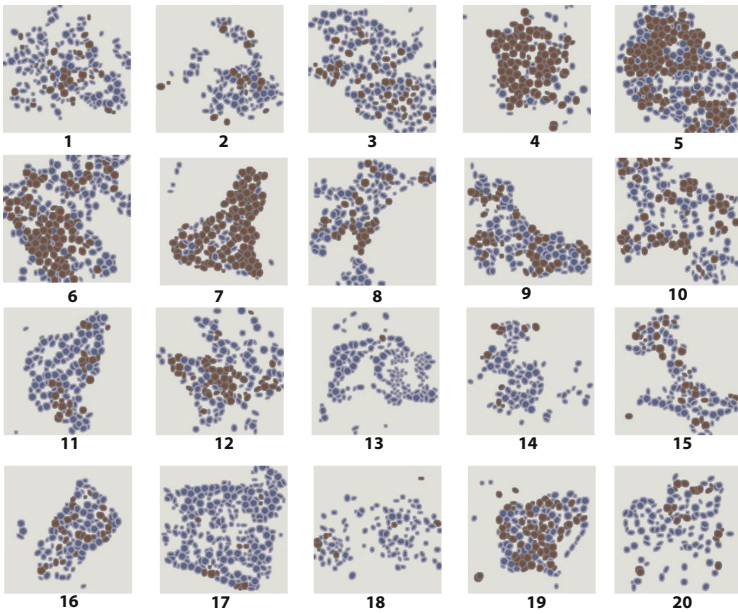


Fig. 4. Segmentation results (Ki-67 positive nuclei are marked with a darker color)

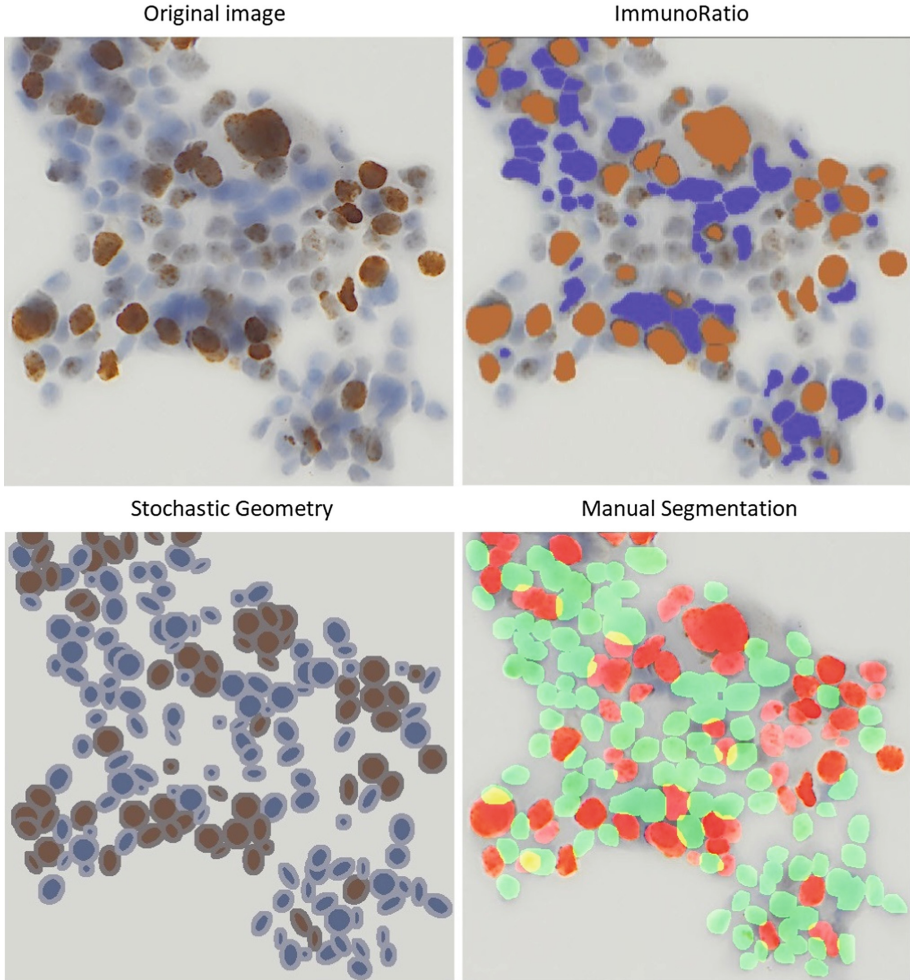


Fig. 5. Segmentation results for image no. 10

stochastic geometry are β and γ . As a result of the experiments, it was found that the best results were obtained for $\ln(\beta) = -600$, and $\ln(\gamma) = -700$.

Segmentation results for all images are presented in Fig. 4. An illustrative example of segmentation results for stochastic geometry, ImmunoRatio, and manual segmentation is presented in Fig. 5. Based on the segmentation results, we were able to estimate the Ki-67 indexes for all test samples. Table 1 summarizes the results obtained for stochastic geometry, ImmunoRatio system and manual segmentation. According to St. Gallen consensus in 2011 and 2013 cut-off points to classify Ki-67 proliferative activity were defined at 14% and 20% levels respectively. We used these cut-off points to classify our results for stochastic geometry, ImmunoRatio, and manual segmentation. Classification results based

Table 1. Results of automatic detection methods in comparison with manual segmentation

Image no.	Stochastic geometry [%]	ImmunoRatio [%]	Ground truth [%]
1	19.66	18.30	19.31
2	17.54	12.40	17.06
3	20.19	23.30	31.42
4	70.50	90.70	62.24
5	48.63	18.00	37.77
6	43.72	28.90	36.94
7	79.47	70.60	65.08
8	27.66	24.00	24.19
9	34.84	25.40	27.20
10	35.47	35.60	36.84
11	17.42	15.00	13.89
12	32.72	25.70	39.02
13	0.64	1.90	1.08
14	9.71	6.50	8.57
15	25.20	31.70	24.50
16	25.44	18.80	21.60
17	5.02	2.90	4.88
18	7.89	16.80	9.02
19	47.86	38.60	45.75
20	17.24	16.20	10.17

Table 2. Confusion matrix (cut-off equal to 14%)

		Reference results	
		Low Ki-67	High Ki-67
Stochastic Geometry	Low Ki-67	4	0
	High Ki-67	2	14
ImmunoRatio	Low Ki-67	3	1
	High Ki-67	3	13

on automatic segmentation were compared with the reference results using confusion matrices (see Tables 2 and 3). Moreover, we computed Cohen's kappa coefficient κ to measure agreement between automatic and manual approaches. Stochastic geometry showed high agreement with the reference data because it obtained $\kappa_{14} = 0.74$ and $\kappa_{20} = 1$ for the 14% and 20% cut-off points respectively. ImmunoRatio got worse results, $\kappa_{14} = 0.47$ and $\kappa_{20} = 0.8$ respectively.

Table 3. Confusion matrix (cut-off equal to 20%)

		Reference results	
		Low Ki-67	High Ki-67
Stochastic Geometry	Low Ki-67	8	0
	High Ki-67	0	12
ImmunoRatio	Low Ki-67	8	2
	High Ki-67	0	10

4 Conclusions

We proposed a novel method based on stochastic geometry for estimating Ki-67 index. The preliminary results of the experiments carried out are satisfactory. We showed that this method is more accurate than the approach used in the ImmunoRatio system. Unfortunately, the number of test images is relatively small, thus the statistical significance of the results is not satisfactory. To conduct more reliable tests, a large number of ground truth images must be provided. Unfortunately, manual counting is tedious and requires the involvement of highly qualified medical personnel. For this reason, we decided that the process of building the reference database will be continuous and experimental results will be updated with the expansion of this database.

The segmentation method proposed in the work belongs to a group of methods with high demand for computing power. However, we managed to implement the presented algorithm with intensive use of convolution operation. This allowed us to run our method using parallel computing on GPU. Moreover, we have shown in Eq. 9 that due to iterative nature of the proposed segmentation method only a small part of the prospective ellipses must be evaluated in the consecutive steps of the algorithm. Only in the first step, all prospective ellipses have to be evaluated. Thanks to such manipulations, it was possible to shorten the time needed to process an image containing approximately 250 nuclei up to 2 min.

Future work will concentrate on building more sophisticated models of nuclei based on additional information from input images such as edges and textures [11]. This will allow us to segment the overlapping cell nuclei more accurately.

Acknowledgments. The research was supported by National Science Centre, Poland (2015/17/B/ST7/03704).


References

1. Abubakar, M., et al.: High-throughput automated scoring of Ki67 in breast cancer tissue microarrays from the Breast Cancer Association Consortium. *J. Pathol: Clin. Res.* **2**(3), 138–153 (2016)
2. Baddeley, A.J., van Lieshout, M.N.M.: ICM for object recognition. In: Dodge, Y., Whittaker, J. (eds.) *Computational Statistics*, vol. 2, pp. 271–286. Physica-Verlag HD, Heidelberg (1992). https://doi.org/10.1007/978-3-642-48678-4_34

3. Baddeley, A.J., van Lieshout, M.N.M.: Stochastic geometry models in high-level vision. In: Mardia, K.V., Kanji, G.K. (eds.) *Advances in Applied Statistics, Statistics and Images*, vol. 1, pp. 231–256. Carfax Publishing, Abingdon (1993)
4. Bembenik, R., Jóźwicki, W., Protaziuk, G.: Methods for mining co-location patterns with extended spatial objects. *Int. J. Appl. Math. Comput. Sci.* **27**(4), 681–695 (2017)
5. Cheang, M.C., Chia, S.K., Voduc, D., Leung, S., Snider, J.: Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl. Cancer Inst.* **101**(10), 736–750 (2009)
6. Descombes, X.: *Stochastic Geometry for Image Analysis*. Wiley, Hoboken (2012)
7. Descombes, X.: Multiple objects detection in biological images using a marked point process framework. *Methods* **115**(Suppl. C), 2–8 (2017)
8. Kowal, M., Filipczuk, P.: Nuclei segmentation for computer-aided diagnosis of breast cancer. *Int. J. Appl. Math. Comput. Sci.* **24**(1), 19–31 (2014)
9. van Lieshout, M.N.M.: Markov point processes and their applications in high-level imaging. *Bull. Int. Stat. Inst.* **56**, 559–576 (1995)
10. Markowsky, P., Reith, S., Zuber, T.E., König, R., Rohr, K., Schnörr, C.: Segmentation of cell structures using model-based set covering with iterative reweighting. In: *2017 IEEE 14th International Symposium Biomedical Imaging (ISBI 2017)*, pp. 392–396, April 2017
11. Mazurek, P., Oszutowska-Mazurek, D.: From the slit-island method to the Ising model: analysis of irregular grayscale objects. *Int. J. Appl. Math. Comput. Sci.* **24**(1), 49–63 (2014)
12. Ruifrok, A.C., Johnston, D.A.: Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* **23**(4), 291–299 (2001)
13. Saha, M., Chakraborty, C., Arun, I., Ahmed, R., Chatterjee, S.: An advanced deep learning approach for Ki-67 stained hotspot detection and proliferation rate scoring for prognostic evaluation of breast cancer. *Sci. Rep.* **7**(1), 3213 (2017)
14. Strauss, D.J.: A model for clustering. *Biometrika* **62**(2), 467–475 (1975)
15. Tuominen, V.J., Ruotoistenmäki, S., Viitanen, A., Jumppanen, M., Isola, J.: ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast Cancer Res.* **12**(4), R56 (2010). <https://doi.org/10.1186/bcr2615>
16. Vörös, A., Csörgő, E., Nyári, T., Cserni, G.: An intra- and interobserver reproducibility analysis of the Ki-67 proliferation marker assessment on core biopsies of breast cancer patients and its potential clinical implications. *Pathobiology* **80**, 111–118 (2013). <https://doi.org/10.1159/000343795>
17. Xing, F., Su, H., Neltner, J., Yang, L.: Automatic Ki-67 counting using robust cell detection and online dictionary learning. *IEEE Trans. Biomed. Eng.* **61**(3), 859–870 (2014)
18. Yeo, M.K., Kim, H.E., Kim, S.H., Chae, B.J., Song, B.J., Lee, A.: Clinical usefulness of the free web-based image analysis application ImmunoRatio for assessment of Ki-67 labelling index in breast cancer. *J. Clin. Pathol.* **70**(8), 715–719 (2017). <http://jcp.bmj.com/content/70/8/715>



Detection Methods of Static Microscopic Objects

Libor Hargaš¹ , Zuzana Loncová², Dušan Koniar¹,
František Jablončík¹, and Jozef Volák¹

¹ Department of Mechatronics and Electronics,
Faculty of Electrical Engineering, University of Žilina, Žilina, Slovakia
libor.hargas@fel.uniza.sk

² Division of Bioinformatics, Biocenter, Medical University of Innsbruck,
Innsbruck, Austria

Abstract. The article deals with selected methods of automated detection of microscopic objects in video sequences obtained by high-speed cinematography and light microscopy. The objects of interest are represented by cilia of airways and also artefact generating objects (gas bubbles and erythrocytes). The main idea of this work is to create complex diagnostic tool for evaluation of ciliated epithelium in airways, where the ratio between moving and static cilia helps to search proper diagnosis (confirmation of PCD – primary ciliary dyskinesia). Methods for automated segmentation of static cilia creates a big challenge for image analysis against the dynamic ones due to character and parameters of obtained images. This work is supported by medical specialists from Jessenius Faculty of Medicine in Martin (Slovakia) and proposed tools would fill the gap in the diagnostics in the field of respiratory in Slovakia.

Keywords: Cilia · Image segmentation · Classification

1 Introduction

When dealing with issues of detection methods of microscopic objects in video sequences (i.e. in series of images), the main research is oriented into the field of image processing and segmentation as the main task leads to searching for some objects within an image content. Image segmentation means choosing the specific wanted objects, so called objects of interest, in an image in order to identify them for further analysis [1].

This article deals with lots of various segmentation methods that could be used when separating the content of given specific video sequences: high-speed video recordings of microscopic sample taken from epithelium of human airways (so called respiratory epithelium). Such video sequences capture the movement of microscopic structures which can be found on the surface of cells of this epithelium, called cilia of respiratory epithelium. Moving (beating) cilia have in one's airways an irreplaceable function: these structures assign the movement of mucus and so clean the whole respiratory apparatus and remove undesirable foreign particles out of airways (this is called a mucociliary transport) [2]. Without this mechanism or just when cilia do not

work properly or stopped working, a serious health problem arises. There is no clearance in the airways and so many inflammations occur due to foreign particles from outer environment. Moreover, as the mucus is present in respiratory system but the structures that should move it do not work, it remains stable and may cause even suffocation, especially at new-born infants. Such medical facts emphasise the importance of cilia and answers the question why to study them and investigate their movement (and so their proper function).

Investigated video sequences are in fact high-speed camera records which captures the scene within microscopic section. The section consists of a sample taken from human respiratory epithelium, usually from nasal or tracheal mucosa, which in a short time after removal from organism (approx. 10 min) contains oscillating hair-like structures, called cilia of respiratory epithelium. Its function lies in cleaning of airways, i.e. removal of foreign bodies such as dust particles, viruses, bacteria or allergens. Malfunction or dysfunction of these structures leads to often, chronic inflammations of respiratory system and to other related serious or even life-threatening diseases. The main objective of monitoring the proper function of cilia using the high-speed camera records is to properly diagnose the patients suspicious to suffer from ciliary pathologies (Fig. 1).

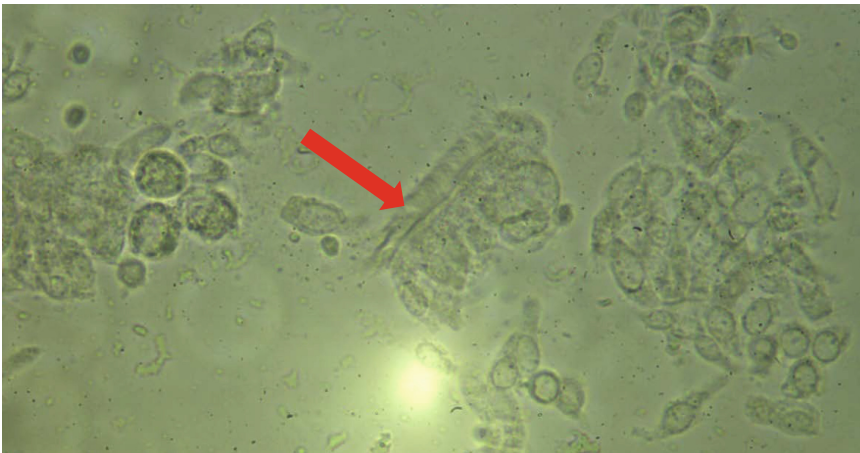


Fig. 1. Example of obtained video sample with ciliary structures in the middle

Cilium is a microscopic structure with length of approx. 6–10 μm and width less than 1 μm . Each cell of respiratory epithelium carries on its surface about 200–300 cilia.

Cilia oscillate synchronously with frequency of 18–30 Hz, while the speed of their beating is referred as CBF (Ciliary Beat Frequency) and represents a crucial factor when determining ciliary pathologies [3, 4]. Movement of cilia functions for sweeping mucus out of human lungs (oral direction) with speed approx. 4–20 mm/min. Beating cycle of each cilium can be divided into two stages: the stage of power stroke (with

effective forward movement) and the stage of recovery stroke (return, backward movement). Power stroke takes only a quarter of beating cycle and means swift movement of cilium which sweeps the mucus, recovery stroke takes three quarters of time needed for one beating cycle and it is a period when cilium stores energy for the next beat.

2 Segmentation Methods Suitable for Static Object

Image operations can be widely classified into four categories (Fig. 2) [5]:

- pre-processing,
- visualization,
- image manipulation,
- and image analysis.

Segmentation is a commonly used operation in pre-processing approaches and an essential operation for most visualization, manipulation and analysis tasks in image processing. Therefore, segmentation is the most crucial as well as the most challenging among all image procedures.

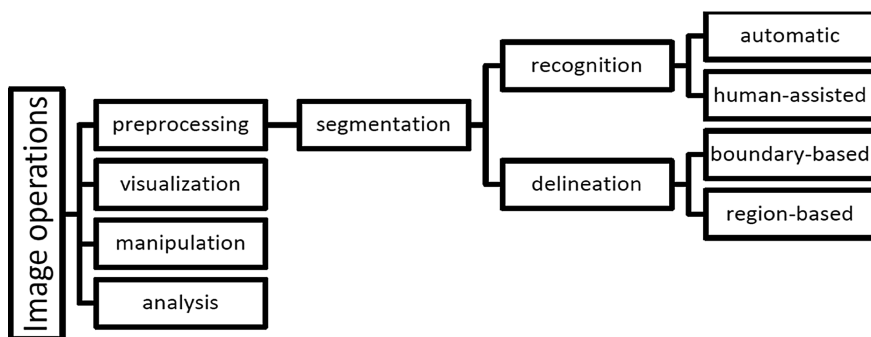


Fig. 2. Classification of image operations

The fact that objects of interest can be within one video either static or dynamic needs to be taken into account right in the beginning when choosing the segmentation approach. Based on the type of sought objects, segmentation approaches can be divided into static methods (i.e. those which are able to find static, non-motile objects in a particular frame of a video) and dynamic ones (i.e. those that identify objects of interests based on their movement through the whole video sequence).

Image segmentation means to categorize an image content into systematized regions or to separate objects of interest from background, i.e. unimportant structures in an image. In general, the purpose of image segmentation is to identify and delineate objects. In case of medical or biomedical images, the objects refer to anatomic organs, tumours, cells, etc. [6–9]. Segmentation can be understood as the partitioning of an

image into non-overlapping regions which are homogenous with respect to some characteristic, such as intensity or texture.

Typically, image segmentation consists of two related tasks:

- recognition
- and delineation of an object.

Recognition means the approximate determining of the objects' location in an image, without determination of its precise specification of the region occupied by the object.

Delineation involves determining the objects' precise spatial extent and composition. Approaches to delineation can be classified as boundary and region based [10].

There are numerous approaches regarding image segmentation techniques in various application and it can be said, that various methods are suitable or certain type of images (with different success) [11]. Some of them are based on simple histogram partitioning, some use various spatial details and others use combined techniques, however, most of them are very sensitive to noise. In particular, segmentation approaches can be classified according to the methodology used in segmentation strategy.

2.1 Basic Thresholding

Thresholding belongs to one of the simplest segmentation techniques based on intensities of particular image pixels. First, a certain value of intensity level is predefined (i.e. threshold), for grayscale images it ranges from 0 to 255. Each pixel of this image is then compared with the threshold: if the value of intensity level of a particular image pixel is greater than selected threshold value, then the new value of this pixel is set to level 255 (i.e. white colour), otherwise it is set to 0 (i.e. black colour) (Fig. 3).

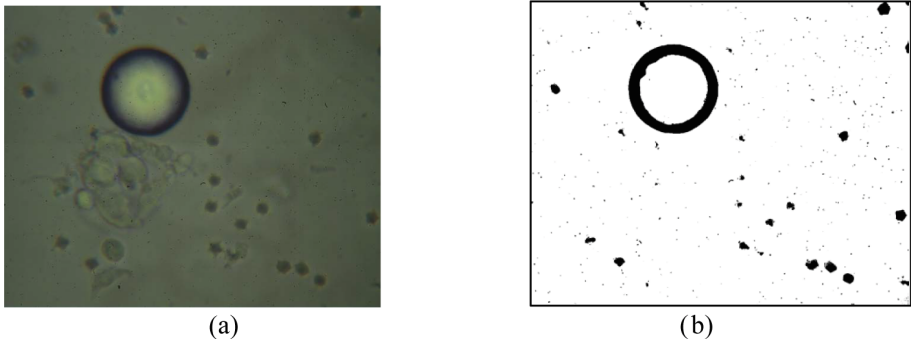


Fig. 3. Example of thresholding of an image containing respiratory epithelium, (a) original image, (b) thresholded image with threshold value $T = 70$

2.2 Region Growing

Principle of region growing method lies in merging neighbouring pixels with similar intensity values so that they form a segmented unit. Usually a two-stage algorithm is used: in the first stage the pairs of quantized pixels are merged into regions called atomic boundaries, and in the next stage the weak boundaries among regions are disbanded [12].

If the assumption is that regions R1 and R2 are two neighbouring regions with perimeter P1 and P2, then regions R1 and R2 are merged if following formula is valid:

$$\frac{D}{\text{MIN}\{P_1, P_2\}} > \varepsilon_1 \quad (1)$$

The method of region growing offers very precise segmentation results of simple scenes with small objects of the same texture, however, it does not work ideally at difficult image scenes (Fig. 4).

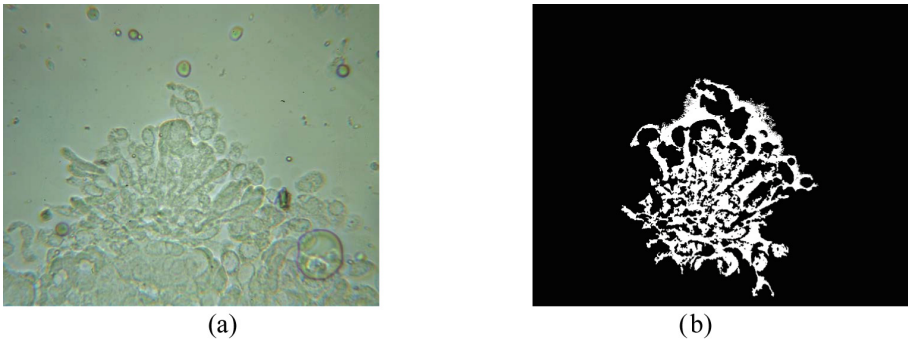


Fig. 4. Example of region growing algorithm applied to an image with respiratory epithelium, (a) original image, (b) segmented image

2.3 Watershed Segmentation

Watershed segmentation means a way how to automatically separate objects in a binary image which are very close or overlapping. This might be especially helpful in cases when the objects need to be counted or unequivocally defined.

Resulting image after watershed segmentation can be displayed either with highlighted boundaries of particular segments or in pseudocolour range. However, as it can be seen in Fig. 5, in case of images with respiratory epithelium this method does not give satisfactory results.

2.4 LBP Method

This method works with image regions as textures, i.e. it tries to identify and separate various image regions based on number of perceived components, using LBP (Local

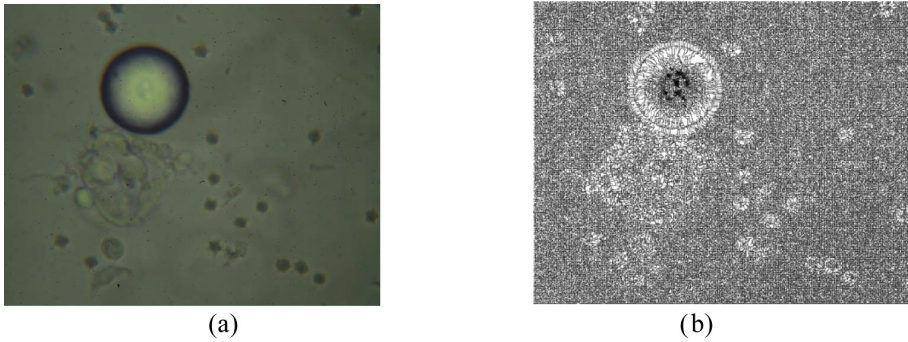


Fig. 5. Example of watershed segmentation applied to an image with respiratory epithelium, (a) original image, (b) segmented image

Binary Pattern) operator [13]. The local binary pattern operator is an image operator which transforms an image into an array or image of integer labels describing small-scale appearance of the image. These labels or their statistics, most commonly the histogram, are then used for further image analysis. The most widely used versions of the operator are designed for grayscale images.

Basic LBP

The basic local binary pattern operator, introduced by Ojala et al. [14] was based on the assumption that texture has locally two complementary aspects, a pattern and its strength. The pixels in this block are thresholded by its centre pixel value, multiplied by powers of two and then summed to obtain a label for the centre pixel. As the neighbourhood consists of 8 pixels, a total of $2^8 = 256$ different labels can be obtained depending on the relative grey values of the centre and the pixels in the neighbourhood. An example of an LBP image and histogram are shown in Fig. 6.

Local Ternary Patterns

A Local Ternary Pattern (LTP) is a generalization of the Local Binary Pattern method, which represents grey level differences between the pixels by a ternary value instead of

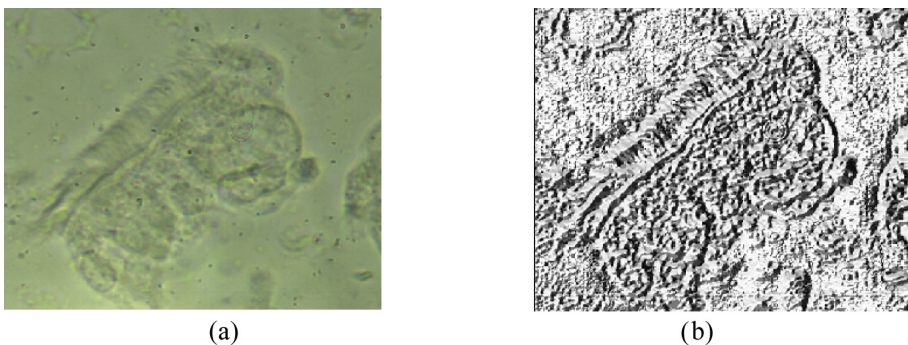


Fig. 6. Example of LBP operator applied to an image, (a) original image, (b) LBP image

a binary one [15]. In order to obtain a more robust descriptor, the difference between the grey value of the pixel x from the grey values of one in its neighbourhood u can assume three values according to a threshold τ :

$$1 \text{ if } u \geq x + \tau; \tag{2}$$

$$-1 \text{ if } u \leq x - \tau; \tag{3}$$

$$\text{else } 0. \tag{4}$$

In this way the descriptor is more discriminant and less sensitive to noise. From a computational point of view, a ternary pattern is split into two binary patterns by considering its positive and negative components.

Local Quinary Patterns

Another extension of LBP algorithm uses five-value encoding, named quinary, in order to obtain a more robust descriptor. In this variant the difference between the grey value of the centre pixel x from the grey values of one of its neighbourhood u can assume 5 values instead of the 2 values as in standard LBP (Fig. 7).

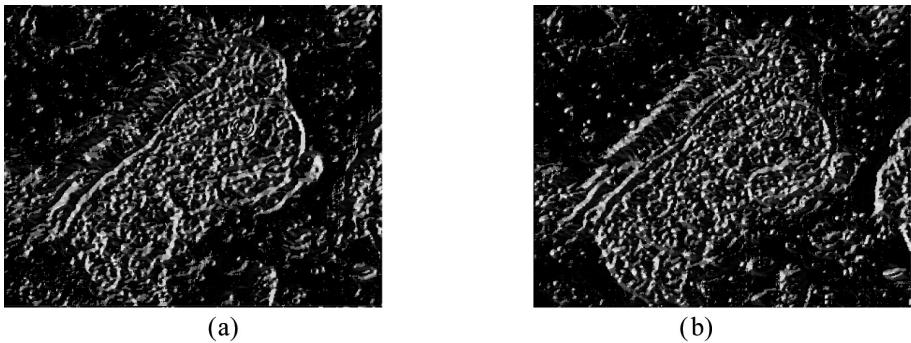


Fig. 7. Examples of LTP method applied to a ciliary image, $\tau = 2$, (a) LTP image – positive LBP and (b) LTP image – negative LBP code

Quinary encoding principles: the difference d between x and u is encoded by 3 values according to two thresholds τ_1 and τ_2 ; $\tau_2 > \tau_1$. The quinary pattern is split into four binary patterns according to the following binary function

$$bc(x), c \in \{-2, -1, 1, 2\} \tag{5}$$

The first binary pattern is obtained by considering $c = 2$, the second considering $c = 1$ and so on for $c = -1$ and $c = -2$. Finally, the histograms computed from the binary patterns are concatenated. The results for local quinary patterns are shown at Fig. 8.

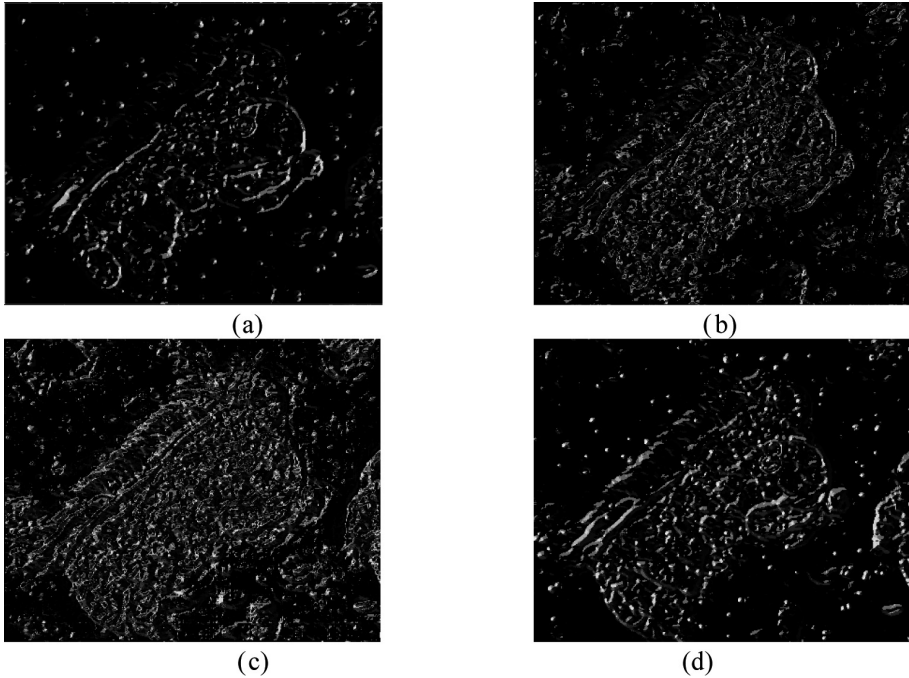


Fig. 8. Examples of LQP method applied to a ciliary image, $\tau_1 = 2$, $\tau_1 = 4$, (a) LQP image for 2 (b) LQP image for 1, (c) LQP image for -1 and (d) LQP image for -2

3 Designed Techniques of Segmentation for Ciliary Images

The proposed method has been tested in several ciliary images with varying contrast, illumination of the microscopic scene and with various proportional representation of objects in the scene: epithelial cells, ciliary areas, artefacts and background. Obtained results have been compared to a set of reference images, i.e. images in which the ciliary areas were labelled by a group of four specialists (doctors who has studied respiratory system and dealt with cilia for more than ten years), and so the precision and effectiveness of proposed algorithm could be evaluated.

Segmentation of any images requires first some pre-processing. This is even more important when dealing with images of respiratory epithelium as they are very specific. First, such images are a bit enhanced using some type of image filtering (smoothing filters). Next difficulty which needs to be eliminated before the segmentation itself, is the fact that investigated images often suffer from low contrast. Traditional (ordinary) histogram equalization was not helpful. This happens due to the fact that investigated (grayscale) images despite their low contrast may contain some pixels with small value of intensity (almost black or 0) as well as some really bright areas (usually as a result of non-homogenous illumination of microscope's scene).

A solution for this difficulty was the usage of some kind of adaptive histogram equalization. As the adaptive histogram equalization (AHE) method suffers from a

tendency to over amplify noise in relatively homogeneous regions of an image, the better results were obtained using contrast limited adaptive histogram equalization (CLAHE) method.

Designed algorithm for detection of ciliary areas can be simply described as the one based on extraction of features from ciliary areas and their consequent classification into defined classes (in general region of interest – ciliary area and image background), using SVM classifier. However, this is a bit more complicated task. In order to achieve the goal of this thesis, the following algorithm consisting of several steps is proposed.

Lots of methods for image segmentation according to some extracted features were tested. The better results (but still not good) came as a result of applying the LBP operator. However, trying different LBP modifications and different topologies of selected LBP neighbourhood for the central pixel gave favourable results with different successfulness for different tested images of respiratory epithelium. All these facts led to a decision to try a combination of various LBP operators for feature extraction – however, not in the way of its direct combining, but in performing all of them separately and then to concatenate them.

In details, together nine different feature vectors have been extracted from selected areas, all of them uniform LBP patterns: LBP for circle neighbourhood topology, LBP for elliptic neighbourhood topology, LBP for spiral neighbourhood topology, LTP (consisting of 2 LBP patterns in fact) and LQP (which consists of 4 LBP patterns in fact). Each of this nine feature vectors consist of 37 LBP uniform values, one non-uniform value, and moreover, one value for calculated entropy for that particular subregion, so all together, the each of proposed feature vectors comprises 39 values (Fig. 9).

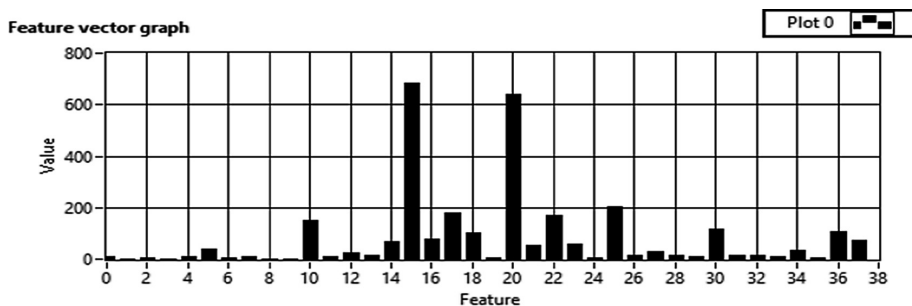


Fig. 9. Example of feature vector for cilia, for LBP with circle neighbourhood and window size of 15 × 15 pixels for performing operator function

The next stage after feature extraction and training is the classification of image content itself (i.e. segmentation in fact). The principle is as follows: a window of selected size moves through the investigated image and in every region computes feature vectors, in the same way as it was defined and used for data training. This currently calculated feature vector (in fact, in this case there are 9 feature vectors) are compared to trained data. A different class can be chosen for each of the 9 vector and so

the finally assigned label of a class is a result of an arbiter: majority decides (5 of 9). As this algorithm assumes original SVM which is a binary classifier, separation of image regions only into two classes can be done in one time.

The original image of respiratory epithelium (not equalized yet) is classified into two classes: epithelium (including everything similar to that) and background, using window size 30×30 . Next, in such a binary image a wider margin is created, by morphological dilations and erosions in several iterations (up to 15). This creates a binary image – mask – of the border between epithelial cells and background in the images – i.e. the place where cilia could be found with the highest probability (as they are in fact structures located on the surface of such cells). In the next step this binary mask is applied to overlay the equalized original image and then another classification is performed, now using window 15×15 (to capture ciliary details) and separating the image content into two classes: cilia and the rest (i.e. background, masked black areas or whatever like that). Finally, small particles are removed from the obtained image and the result is then overlayed with original image to display and label where in the investigated image the ciliary area(s) has (have) been identified.

Individual steps of this proposed algorithm for detection of (static) cilia within an image are illustrated in the following example (Fig. 10).

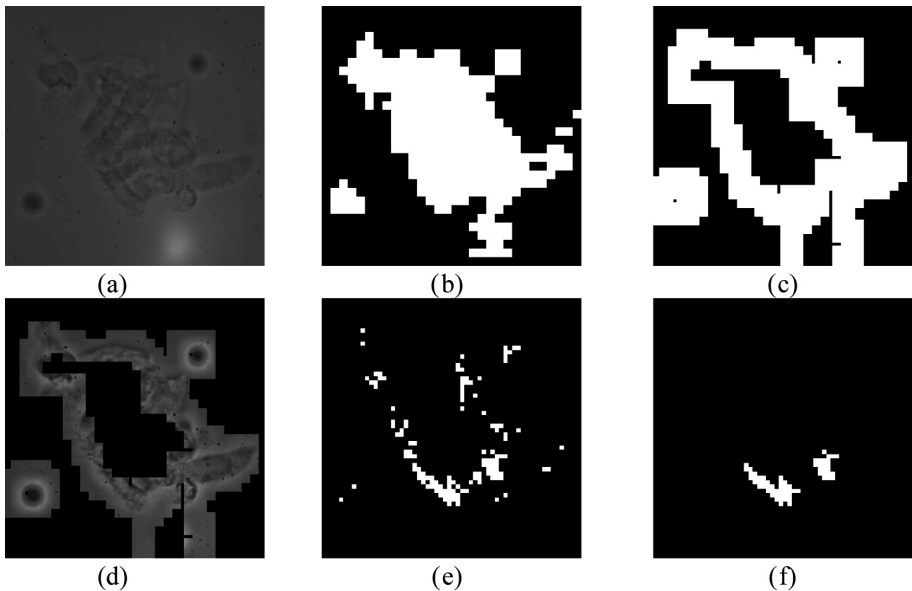


Fig. 10. Example of designed algorithm for detection of static cilia within image of respiratory epithelium (a) original image (not equalized) (b) result after first classification: epithelium and background (c) created margin between background and epithelium (d) overlaying with equalized original image (e) result after second classification: cilia and the rest (f) result after small particles removal

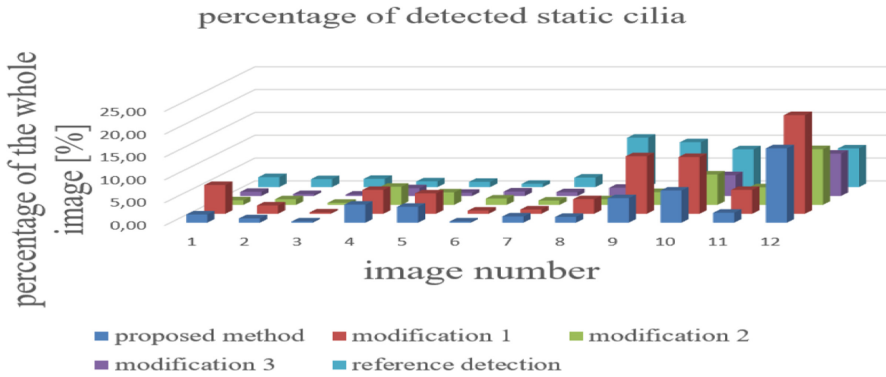


Fig. 11. Graphical representation of detected ciliary areas within 12 tested images of respiratory epithelium, for four tested methods compared to reference detection

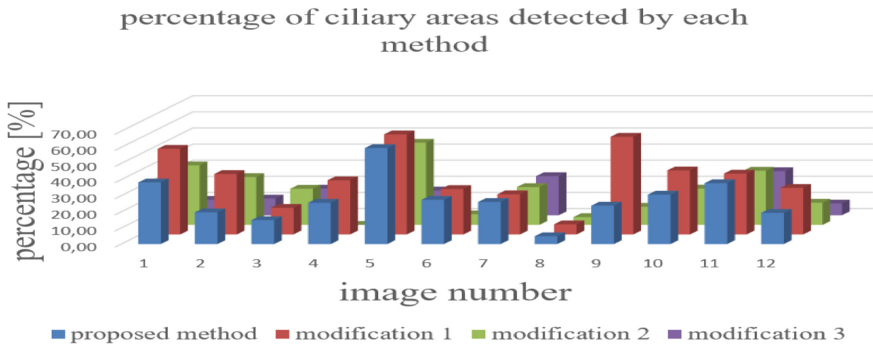


Fig. 12. Graphical representation of comparison of common intersection between each tested method and reference ciliary area

4 Conclusion

This proposed segmentation method has been tested in 12 selected images of respiratory epithelium, representing the samples of various types (e.g. including artefacts, with overlapping epithelial cells, ones with small ciliary areas, ones with large areas, ones with different proportion of background and epithelium, etc.). Obtained results are shown in following graphs (Figs. 11 and 12) and compared to reference images by counting the overlapping area. Assessment criteria such as precision, recall or accuracy were also evaluated.

Figure 11 illustrates the static ciliary areas detected within investigated images by four different methods, compared to reference segmentation of ciliary areas (given in percentage of the area of the whole image). This provides an information about how many cilia are actually present in the obtained sample (reference detection) and compares this with proposed methods. In details, in this graph it can be seen in which

cases the tested methods detected even more ciliary areas than actually present in certain image, i.e. this indicates that also false-positive detections occurred. As this graphical representation evaluates only the quantity of detected areas of interest, the next one (Fig. 12) assumes also its location. It evaluates the common intersection of results from tested methods with the reference, i.e. it shows the percentage of correctly detected (and truly marked) ciliary areas.

The results of detected areas show that this method for detection of static areas identifies significantly less area of overall image. This could indicate that such a proposed algorithm focuses on image details and that is the reason why it does not detect large image areas and thanks to this it is able to find some static, not moving cilia. Evaluation of common intersection between each tested method and reference segmentation show that the worst results were obtained by modification 3 of proposed method, i.e. starting the algorithm's operation in equalized image. The best results, i.e. the highest values of intersections with reference areas (true ciliary areas) in most images were obtained using the modification 1 of proposed algorithm, i.e. starting in original, not equalized image and with window sized 15×15 pixels. These results were quite similar to those obtained by starting in original image without equalization and window sized 30×30 pixels. As the modification 2 of proposed algorithm (starting with equalized image, window 15×15) shows the second worst results of the four tested methods, it can be concluded that usage of equalization right in the beginning is not suitable and leads to some wrong classifications and so to inappropriate segmentation of image content.

Acknowledgement. Authors of this paper wish to kindly thank to all supporting bodies, especially to grant APVV-15-0462: Research on sophisticated methods for analyzing the dynamic properties of respiratory epithelium's microscopic elements.


References

1. Russ, J.C.: The Image Processing Handbook, 6th edn. CRC Press, Boca Raton (2011). ISBN 978-1-4398-4063-4
2. Trojan, S., et al.: Lékařská fyziologie, 4th edn. Grada Publishing, Praha (2003). ISBN 80-247-0512-5
3. Nečas, E., Šulc, K., Vokurka, M.: Patologická fyziologie orgánových systémů. Nakladatelství Karolinum, Praha (2006). ISBN 80-246-0675-5
4. Silbernagl, S., Despopoulos, A.: Atlas fyziologie člověka, 8th edn. Grada Publishing, Praha (2016). ISBN 978-80-247-4271-7
5. Pei-Gee, P.H. (ed.): Image Segmantation. InTech, Rijeka (2014). ISBN 978-953-307-228-9
6. Haralick, R.M., Shapiro, L.G.: Image segmentation techniques. *Comput. Vis. Graph. Image Process.* **29**, 100–132 (1985). ISSN 0734-189X
7. Frankovsky, P., Ostertag, O., Trebuna, F., Ostertagova, E., Kelemen, M.: Methodology of contact stress analysis of gearwheel by means of experimental photoelasticity. *Appl. Opt.* **55** (18), 4856–4864 (2016). ISSN 1539-4522
8. Parker, J.R.: Algorithms for Image Processing and Computer Vision, 2nd edn. Wiley Publishing, Hoboken (2010). ISBN 978-0-470-64385-3

9. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Pearson, London (2008). ISBN 978-0131687288
10. Umbaugh, S.E.: Computer Imaging: Digital Image Analysis and Processing. CRC Press, Boca Raton (2000). ISBN 0-8493-2919-1
11. Mikulova, Z., Duchon, F., Dekan, M., Babinec, A.: Localization of mobile robot using visual system. *Int. J. Adv. Robot. Syst.* **14**(5). Article No. 1729881417736085. ISSN 1729-8814
12. Pratt, W.K.: Digital Image Processing: PIKS Inside, 3rd edn. Wiley, New York (2001). ISBN 0-471-22132-5
13. Pietikainen, M., Hadid, A., Zhao, G.: Computer Vision Using Local Binary Patterns. *Computer Imaging and Vision*, vol. 40. Springer, London (2011). <https://doi.org/10.1007/978-0-85729-748-8>. ISBN 978-0-85729-748-8
14. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.* **29**, 51–59 (1996). ISSN 0031-3203
15. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. In: Zhou, S.K., Zhao, W., Tang, X., Gong, S. (eds.) AMFG 2007. LNCS, vol. 4778, pp. 168–182. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75690-3_13. ISBN 978-3-540-75690-3



Parkinson's Disease Database Analysis of Stereotactic Coordinates Related to Clinical Outcomes

Francisco Estella¹ , Esther Suarez², Beatriz Lozano², Elena Santamarta², Antonio Saiz², Fernando Rojas¹, Ignacio Rojas¹, and Fernando Seijo²

¹ University of Granada, Granada, Spain

francisco.estellap@gmail.com

² Hospital Universitario Central de Asturias, Oviedo, Spain

Abstract. Parkinson's Disease is one of the leading movement disorder diseases. It is the fourth most common neurological disease, after migraine, stroke and epilepsy. The motor symptoms of the disease significantly impair daily living and quality of life and exact a high burden on both patients and their caregivers. Deep Brain Stimulation is a proven therapy for this disease, getting positive outcomes while reducing medication. In this paper, stereotactic system used for Deep Brain Stimulation (DBS) procedures will be described. Different planning methods will be observed and compared to the gold standard normally used, neurophysiological coordinates recorded intra-operatively. MRI, CT scan and direct calculation of stereotactic coordinates will be compared and group in three different groups, according to DBS therapy outcomes: "very good DBS therapy", "good DBS therapy" and "not major improvement".

Database of 72 DBS electrodes implanted in Parkinson's Disease patients will be studied. Most potentially beneficial ranges of deviation within planning and neurophysiological coordinates from the operating room will be assessed, in order to provide neurosurgeons with more landmarks in order to achieve the best outcomes within a millimetric technique.

We could confirm three main highlights out of this study: the neurophysiological length of the Subthalamic Nucleus does not play a major role in outcomes while being within normal range; CT scan calculations were the most accurate; direct calculations should not be used as major deviations were observed.

Keywords: Parkinson's disease · Deep brain stimulation
Movement disorders · Decision tree · Stereotactic coordinates

1 Introduction

Parkinson's disease is a chronic, neurodegenerative disease whose chief symptoms are motor difficulties such as tremor, rigidity, bradykinesia and postural instability.

It is the fourth most common neurological disease, after migraine, stroke and epilepsy.

The motor symptoms of the disease significantly impair daily living and quality of life and exact a high burden on both patients and their caregivers.

About 0.3% of the population, or 1.2 million Europeans, are currently suffering from Parkinson's disease.

As the population ages, the prevalence of Parkinson's disease will increase [1, 2].

Demographic studies show that elderly individuals' proportion will increase rapidly in the following years. The general evolution of medicine, together with a much better quality of life, will be translated in an increase of population age. Also the decrease in birth rates and child mortality, lead us also to an unavoidable increase in the average age of world population [3].

In Spain, according the National Institute of Statistics, in 2020 19.7% of population will be more than 65 year-old, and in 2050 this percentage will get to 31.2%, being the group older than 75 year-old the one increasing the most, from 7.6% in 2002 to 16.8% in 2050. In 2050 Spain will also be the European country with the highest rate of people older than 65 year-old; and also in that year it will be the third one with the highest percentage of people older than 80 year-old (12.8% of the population). Only Italy (14.1%) and Germany (13.6%), will have higher rates.

Even more interesting is the aging of population in developing countries, which would go from a poverty to an elderly population stage, without going through an stable and wealthy state.

This increase in age for the population will be associated to an increase of chronic diseases and disabilities, both in the physical and also the psychiatric and social areas. As discussed, diseases as Parkinson's Disease will increase their prevalence.

Nowadays 1.2 million Europeans are affected by Parkinson's Disease, with main countries being observed in the following figure (see Figs. 1 and 2):



Fig. 1. PD prevalence per country in Europe

Deep brain stimulation (DBS) is a surgical procedure used to treat the motor symptoms of Parkinson's disease.

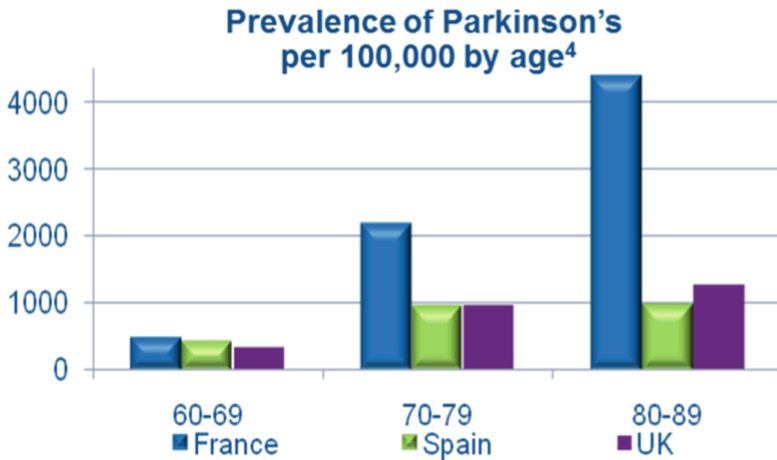


Fig. 2. Prevalence of PD per 100,000 by age

DBS uses a surgically implanted pulse generator to deliver electrical stimulation to the brain through leads with contact electrodes in their distal end. This is performed using Stereotactic techniques.

The electrode leads are implanted in the brain using MRI or CT scan to identify the correct area of the brain (thus defining coordinates), to target with stimulation, usually the subthalamic nucleus, thalamus or globus pallidus.

Once implanted, the leads are connected to the implantable pulse generator (IPG), generally implanted beneath the clavicle, with an extension wire which runs under the skin [5, 6].

Main study objectives are:

- Analyze a database of 72 implanted electrodes for Deep Brain Stimulation using stereotactic systems, to treat Parkinson's Disease. Define if a surgery of this kind would be successful in advanced through different patterns of coordinates deviation within CT scans or MRI, in relation to the "Gold Standard" from the operating room: neurophysiological coordinates.
- Specify a maximum deviation range calculated during the planning stage of the surgery, to assure a successful therapy after implantation.

2 Materials and Methods

Parkinson's Disease patients implanted with Deep Brain Stimulation systems (electrodes and implantable pulse generator, IPG) following standard DBS surgical procedure for planning and implantation (see Fig. 3).

Food and Drug Administration (FDA), approved DBS as a treatment for Essential Tremor in 1997, for Parkinson's Disease in 2002 and for Dystonia in 2003. DBS is also used for chronic pain and has been used to treat or psychiatric disorders, such as depression or OCD.

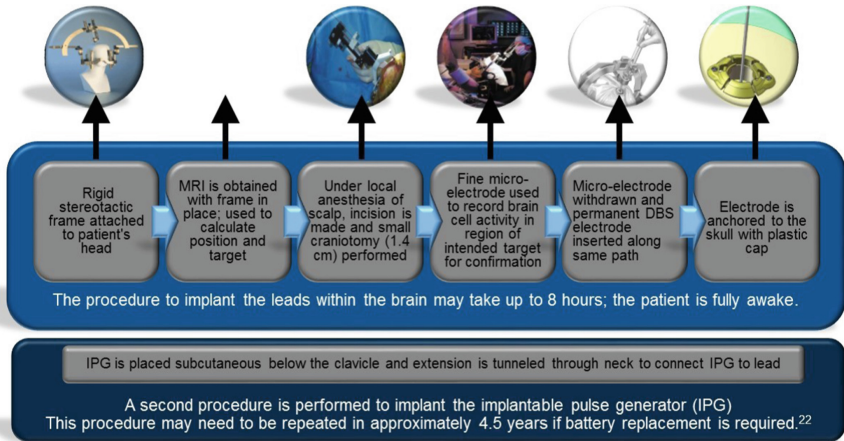


Fig. 3. DBS surgical procedure

The most common brain area used nowadays for Parkinson's Disease treatment with DBS is subthalamic nucleus, which is a brain nucleus of tiny size (3–4 mm high \times 2–3 mm width).

Neighbor structures stimulation could lead to undesired side effects when stimulating, thus bad clinical results. Precision in electrodes location is critical during the surgical procedure. This is why stereotactic coordinates system is used, in order to place the target structure and place the electrode in the right spot. These coordinates are applied using a stereotactic frame located in patient's head (see image).

Medical images acquisition is then performed (CT scan and MRI), so anterior and posterior commissures could be identified. Afterwards the target is identified using them, and this "target coordinates" (see Fig. 4) are then translated and placed in the frame to perform the external adjustment in the frame, in order to be able to place the electrodes deep in the brain correctly.

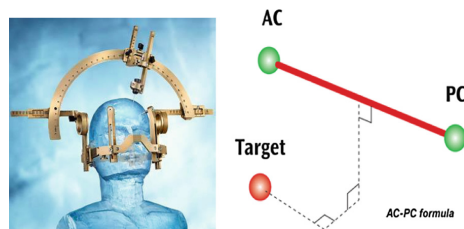


Fig. 4. Stereotactic frame system and AC-PC coordinates to target

Seventy-two electrodes implanted in Hospital Central de Asturias (HUCA); by Dr. Fernando Seijo's team using planning through CT and MRI scans taken pre-operatively.

Stereotactic coordinates are calculated by direct targeting of the subthalamic nucleus and by manual calculation from Anterior and Posterior Commissures (AC-PC), landmarks.

Neurophysiological coordinates and target length (subthalamic nucleus length) calculated intraoperatively with a Micro Electrode Recording (MER) system, targeting first theoretical coordinates (calculated during planning phase); and then shifting accordingly to neurophysiological and clinical response (Fig. 5).

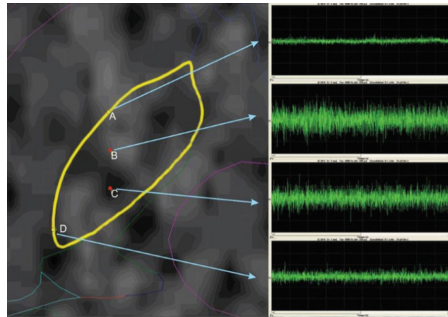


Fig. 5. Subthalamic nucleus MER recording

Patients were then treated with the stable combination within Deep Brain Stimulation and adjusted (and minimized) medication, Levodopa.

UPDRS III scale, which assess motor symptoms and it's the most widely performed to validate DBS results, was used to assess the outcome for the therapy and baseline level of disease.

33% of patients were classified as very successful treatment; 37% as successful treatment and 14% as not successful treatment. This range was assessed though a meta-analysis of different publications up to date (see Fig. 6) (Table 1).

As we could observe, maximum improvement average was around 55% in UPDRS III; however as recent studies are already showing around 60% improvement due to new technologies [16], including a trial performed in the center of reference studied in this article. Thus, the ranges were defined as a very successful therapy (as significantly better than previous studies), for patients treated with DBS with more than 60% improvement in UPDRS III scale.

Other scales such as UPDRS II or medication reduction were also studied, however as UPDRS III scales measuring motor symptoms is the most widely used, the results of this paper are focused on those three sub-group based on UPDRS III percentage of improvement (Table 2).

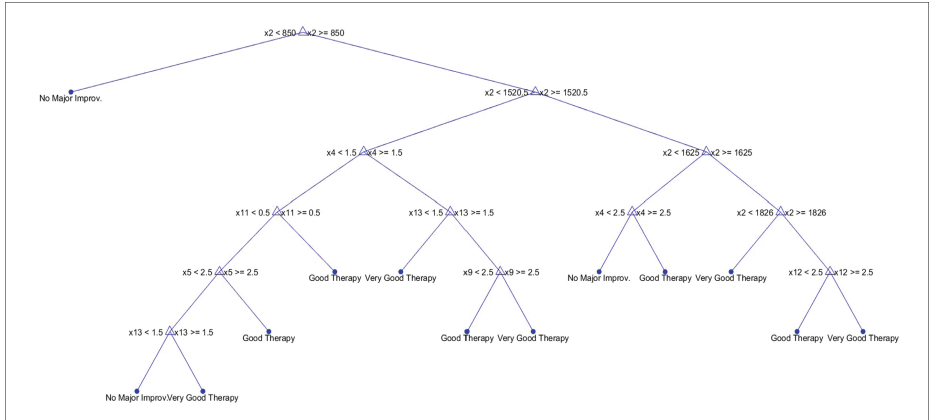


Fig. 6. Decision tree classification including medication as variable of study

Table 1. Comparison groups depending on clinical outcomes due to DBS

Very successful treatment	Successful treatment	Not successful treatment
>60% UPDRS III improvement (vs. pre-op)	[35,60]% UPDRS III improvement (vs. pre-op)	<35% UPDRS III improvement (vs. pre-op)

Table 2. Meta-analysis of previous clinical trial of DBS treatment for PD patients

Primary author	No. of subjects implanted	Study Design	Follow up	Claim
Deuschl 2006 ⁷	78	Randomized pairs trial	6 mos.	• 41% improvement in UPDRS III scores
Follett 2010 ⁸	299	Multi-center, randomized, blinded	24 mos.	• 25.3% improvement in UPDRS III scores
Fraix 2006 ⁹	95	Prospective, multi-center	12 mos.	• 57% improvement in UPDRS III scores
Rodriguez-Oroz 2005 ¹⁰	69	Blinded, multi-center study	3-4 yrs.	• 50% improvement in UPDRS III with STN and 39% improvement with Gpi at 3-4 yrs.
Gervais-Bernard 2009 ¹¹	42	Prospective, single-center	5 yrs.	• 55% improvement in UPDRS III scores
Lefaucheur 2008 ¹²	54	Single-center study	12 mos.	• 56% improvement in UPDRS III scores without medication, 73% improvement with medication
Moro 2010 ¹³	51	Non-randomized, prospect, blinded, multi-center study	5-6 yrs.	• 20% (GPI) to 45% (STN) improvement in UPDRS III scores
Tir 2007 ¹⁴	103	Prospective, single-center study	12 mos.	• 43% improvement in UPDRS III scores
Zangaglia 2009 ¹⁵	32	Prospective, naturalistic controlled study	3 yrs.	• 26% improvement in UPDRS III scores, compared with 15% decline in non-DBS cohort

3 Results

Subthalamic nucleus length recorded through MER system was not of significant value to define better clinical outcomes. Length was within similar range in all three controlled groups ([4.36, 4.50] mm). (see Figs. 6 and 7).

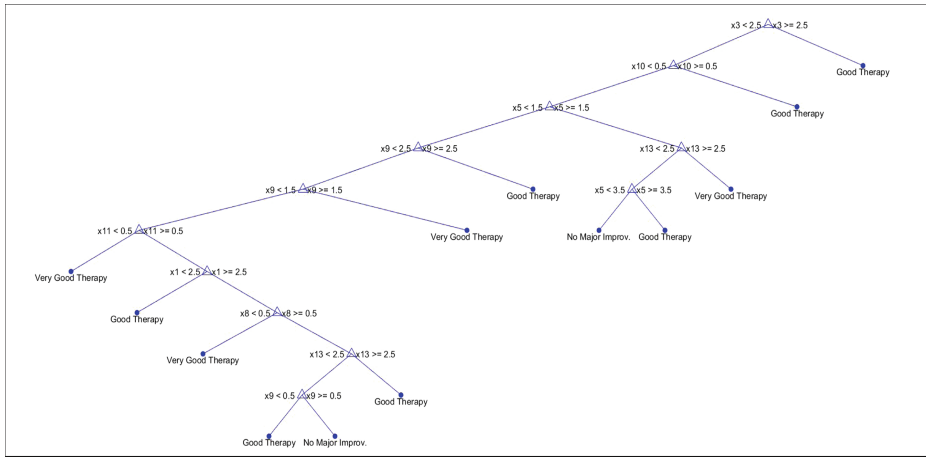


Fig. 7. Decision tree classification without including medication as variable of study

Direct target coordinates calculation shows the highest difference with Gold Standard, getting to major deviation in the Z-Axis (dev > 2.3 mm).

CT-Scan and MRI-scan calculated coordinates show similar behavior in terms of comparison with Gold Standard group for X and Y-axis. However, MRI-scan coordinates differ significantly more than CT-scan coordinates when comparing Z-axis. (see Figs. 6 and 7).

Deviation average compared to “gold standard” neurophysiological coordinates showed no significant value to determine better outcomes, as the three control groups showed similar results for deviations in each coordinate and modality. (see Figs. 6 and 7).

Variables studied and their corresponding classification variable assigned (from X1 to X15), were as followed (Table 3):

Decision tree classification analysis was performed using these 15 clinical variables as input. Classification output was obtained after training the classification method with the different subject’s sub-groups (and 15 clinical variables per subject).

Results were studied following two different methodologies, taking into account initial medication or without taking it into account, getting to the main highlights below:

Table 3. Clinical variables studied with their corresponding classification variable assigned for the study.

Variable X1	Variable X2	Variable X3	Variable X4	Variable X5	Variable X6	Variable X7	
Neurophysiological STN length	Ldopa medication (pre-op) (mg)	Dev. Coordinates Xct-Xnph	Dev. Coordinates Xmri-Xnph	Dev. Coordinates Xdir-Xnph	Dev. Coordinates Xcontrol-Xnph	Dev. Coordinates Yct-Ynph	
Variable X8	Variable X9	Variable X10	Variable X11	Variable X12	Variable X13	Variable X14	Variable X15
Dev. Coordinates Ymri-Ynph	Dev. Coordinates Ydir-Ynph	Dev. Coordinates Ycontrol-Ynph	Dev. Coordinates Zct-Znph	Dev. Coordinates Zmri-Znph	Dev. Coordinates Zdir-Znph	Dev. Coordinates Zcontrol-Znph	L-dopa medication (post-op)(mg) stim ON
		* Variable "ct" * Variable "mri" * Variable "dir" * Variable "control" * Variable "nph"		CT scan image based coordinate mri scan image based coordinate CT scan image based coordinate post-op CT scan image based coordinate "gold standard", neurophysiological recording intra-op based coordinate			

3.1 Variables Classification Accounting Medication

Medication level is a key indicator for DBS therapy success (“not major improvement” for patients with meds (Ldopa) pre-op < 850 mg.

This makes sense initially as most of the eligible patients for DBS therapy are usually pretty advanced PD patients with already high dose of medication. However, with this study we could observe that in cases when anyways DBS is performed, and in which patients are not yet taking high doses of medication, DBS is not providing major improvement in the therapy. We could also observe that the higher the pre-op medication level was, the better results in percentage of improvement in UPDRS III were observed.

We could also observe that the deviation within MRI coordinates and Neurophysiological coordinates when comparing X-axis was providing better outcomes when deviation was higher. This means that when trying to find the subthalamic nucleus intra-operatively, neurosurgeons should take into account that correct deviation should be taken in that X-axis instead of Y-axis.

3.2 Variables Classification Without Medication

As we defined previously, direct coordinates are the ones adding higher deviation compared to “gold standard” neurophysiological coordinates, so we could assess these should not be taking into account primarily when compared to MRI and CT scan coordinates when analysis results.

This leads us, when observing the decision tree classification performed for this chapter, extracting pre-op medication from the variables studied, to define that “very good outcomes” are observed when X-axis CT scan deviation from neurophysiological coordinates is minor to 2.5 mm and also Z-axis deviation with CT scan compared to intra-operative Z-axis is almost zero (<0.5 mm).

Together with the previous analysis using also the pre-op medication indicator, we could observe that potentially, a deviation of >1.5 mm from the MRI coordinates calculated should be taken into account when trying to find the target intra-operatively; but also that neurosurgeons should not deviate further to 2.5 mm from the CT scan X-axis coordinates calculated during planning step.

4 Discussion

Important to note that neurophysiological length of the subthalamic nucleus recorded, does not affect clinical outcomes while being within range (>4 mm).

After analyzing the results, we can confirm CT-scan calculations as the most accurate ones when planning for Deep Brain Stimulation.

Direct calculation should be voided as major deviation in Z-axis was observed, which could be even dangerous during surgical procedure, leading to hemorrhages or undesired stimulation.

Further analysis is needed to assess a specific landmark as key indicator for successful therapy. This is already in progress as Decision Trees and other classification methods are being used to define the key differentiator for a very successful Deep Brain Stimulation Therapy.

5 Conclusion

It seems clear that as technology advances, the need of algorithms to optimize the time of the clinical team is becoming more important. Pre-operative algorithms will lead to optimization for operating room timings, and will allow functional units using Deep Brain Stimulation for a better understanding and estimation of potential positive outcomes for many patients.

On the other hand, we cannot forget the variability within patients; thus being impossible to assess a general and standard methodology for all treated cases.

This study shows promising results in classification for Deep Brain Stimulation cases; allowing physicians to choose CT scan images as initial method when planning coordinates manually for Deep Brain Stimulation. This is increasingly changing as automatic fusion and planning algorithms are being introduced to calculate planning using fewer calculations.

Subthalamic nucleus length recorded was usually believed as an indicator for a better trajectory; however, the opposite has been proven with these results, showing less weight from that variable when getting better clinical outcomes.


More research and development on pre-operative algorithms to potentially assess Deep Brain Stimulation outcomes are needed, but in any case, this study demonstrates that CT scan images are the most accurate ones.

References

1. Parkinson's Disease Foundation. Primary Motor Symptoms. http://www.pdf.org/symptoms_primary. Accessed 25 July 2014
2. European Brain Council. Parkinson's Disease Fact Sheet 2011. <http://www.europeanbraincouncil.org/pdfs/Documents/Parkinson's%20fact%20sheet%20July%202011.pdf>. Accessed 25 July 2014
3. Lasprilla, J.C.A., Fernández Guinea, S., Ardila, A.: Las demencias. ASpectos clínicos, neuropsicológicos y tratamiento. Manual Moderno
4. Lutz, W.: European Demographic Data Sheet 2006 (Vienna and Washington, DC: Vienna Institute of Demography, International Institute for Applied Systems Analysis, and Population Reference Bureau) (2006)
5. Dillon, A.: Deep brain stimulation for Parkinson's disease, National Institute for Clinical Excellence (IPG019), November 2003
6. Fisman, G.K., Herzog, J., Fisman, D., Lang, A.E., Deuschl, G.: Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov. Disord.* **21**(Suppl. 14), S290–S304 (2006)
7. Deuschl, G., Schade-Brittinger, C., et al.: A randomized trial of deep-brain stimulation for Parkinson's disease. *N. Engl. J. Med.* **355**, 896–908 (2006)
8. Follett, K.A., Weaver, F.M., Stern, M., et al.: Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *N. Engl. J. Med.* **362**, 2077–2091 (2010)
9. Fraix, V., Houeto, J.L., Lagrange, C., et al.: Clinical and economic results of bilateral subthalamic nucleus stimulation in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **77**, 443–449 (2006)
10. Rodriguez-Oroz, M.C.: Bilateral deep brain stimulation in Parkinson's disease: a multicentre study with 4 years follow-up. *Brain* **128**, 2240–2249 (2005)
11. Gervais-Bernard, H., Xie-Brustolin, J., Mertens, P., et al.: Bilateral subthalamic nucleus stimulation in advanced Parkinson's disease: five year follow-up. *J. Neurol.* **256**, 225–233 (2009)
12. Lefaucheur, J.P., Gurruchaga, J.M., Pollin, B., et al.: Outcome of bilateral subthalamic nucleus stimulation in the treatment of Parkinson's disease: correlation with intra-operative multi-unit recordings but not with the type of anaesthesia. *Eur. Neurol.* **60**, 186–199 (2008)
13. Moro, E., Lozano, A., Pollak, P., et al.: Long-term results of a multicenter study on subthalamic and pallidal stimulation in Parkinson's disease. *Mov. Disord.* **25**(5), 578–586 (2010)
14. Tir, M.: Exhaustive, one-year follow-up of subthalamic nucleus deep brain stimulation in a large, single-center cohort of Parkinson's patients. *Neurosurgery* **61**, 297–305 (2007)
15. Zangaglia, R., et al.: Deep brain stimulation and cognitive functions in Parkinson's disease: a three-year controlled study. *Mov. Disord.* **24**(11), 1621–1628 (2009)
16. Timmermann, L., Seijo, F., et al.: Multiple-source current steering in subthalamic nucleus deep brain stimulation for Parkinson's disease (the VANTAGE study): a non-randomised, prospective, multicentre, open-label study. *Lancet Neurol.* **14**(7), 693–701 (2015)



Quantitative Ultrasound of Tumor Surrounding Tissue for Enhancement of Breast Cancer Diagnosis

Ziemowit Klimonda¹(✉) , Katarzyna Dobruch-Sobczak^{1,2},
Hanna Piotrkowska-Wróblewska¹, Piotr Karwat¹, and Jerzy Litniewski¹

¹ Institute of Fundamental Technological Research,
PAS, Pawińskiego 5b, 02-106 Warsaw, Poland
zklim@ippt.pan.pl

² Maria Skłodowska-Curie Memorial Cancer Centre and Institute of Oncology,
Warsaw, Poland

Abstract. Breast cancer is one of the leading causes of cancer-related death in female patients. The quantitative ultrasound techniques being developed recently provide useful information facilitating the classification of tumors as malignant or benign. Quantitative parameters are typically determined on the basis of signals scattered within the tumor. The present paper demonstrates the utility of quantitative data estimated based on signal backscatter in the tissue surrounding the tumor. Two quantitative parameters, weighted entropy and Nakagami shape parameter were calculated from the backscatter signal envelope. The ROC curves and the AUC parameter values were used to assess their ability to classify neoplastic lesions. Results indicate that data from tissue surrounding the tumor may characterize it better than data from within the tumor. AUC values were on average 18% higher for parameters calculated from data collected from the tissue surrounding the lesion than from the data from the lesion itself.

Keywords: Quantitative ultrasound · Tissue characterisation
Tumor classification

1 Introduction

Breast cancer is one of the leading causes of cancer-related death in female patients worldwide [1], and the second most common in Poland [2]. Methods are sought to enhance the effectiveness of early diagnosis of breast diseases and the accuracy of referral for biopsy. Accurate early diagnosis of breast tumors is important both for treatment selection, and for predicting patients' response to therapy. Currently, two techniques are most commonly used in breast diagnosis: X-ray imaging (mammography) and ultrasound imaging. The former uses ionizing radiation and typically has lower sensitivity in breasts with a predominance of glandular tissue [3, 4]. In ultrasound imaging, image quality and interpretation

results are highly dependent on equipment quality, machine settings adjusted by the operator, and the physician's experience.

Tumor classification methods using B-mode ultrasound images are often based on an analysis of the tumor's anatomical features shown, e.g. shape, borders, echostructure, and echogenicity. The BI-RADS scale is used, providing standardized diagnostic assessment and guidelines for further management [5]. In another approach, quantitative ultrasound (QUS) techniques are used to attempt an estimation of the physical properties of tissue or scatter statistics by analyzing received ultrasound radio-frequency (RF) signals.

A broad spectrum of methods is available. For instance, the physical properties of tissue (attenuation coefficient and backscattering coefficient) were proposed by D'Astous and Foster [6] and Nam et al. [7] for assessing various breast tissue types. The former team presented findings demonstrating that the attenuation coefficient determined for uniform areas of ductal carcinoma was higher than in fat tissue, but lower than in fibrous and parenchymal tissues. In turn, backscattering coefficient values were similar for ductal carcinoma and fat and considerably lower than in the parenchyma. Nam et al. reported attenuation coefficients for carcinoma that were approximately 20% higher than for fibroadenoma. Signal spectrum analysis allows for determining parameters such as spectral slope (SS), spectral intercept (SI), and midband fit (MBF). Lizzi et al. [8] demonstrated a direct correlation between these parameters and tissue microstructure. Spectral slope is inversely proportional to the size of scatterers; SI is determined by scatterer size, spatial density, and the acoustic impedance difference between the scatterers and background; and midband fit is correlated with integrated backscattering coefficient (IBC). Tissue parameters determined using texture parameters of ultrasound [9] and parametric images [10], and signal phase [11] have also been used for breast tumor assessment. An extensive review of classification methods was presented by Cheng et al. [12].

The scattering properties of the medium are often assessed by modeling the scattered signal envelope statistic using the probability density function (PDF). PDF shape parameters have been used to classify breast lesions. One very universal distribution is the homodyned K distribution [13–17]. Trop et al. [16] demonstrated that a combination of this distributions parameters with BI-RADS allows for reducing the number of biopsies of focal breast lesions by 25%, while maintaining a 100% sensitivity. The Nakagami distribution is a simplified form of homodyned K distribution [18], and its shape parameter value depends both on the spatial distribution of scatterers, and on their number and homogeneity. The simple formulation and ease of shape parameter calculation made the Nakagami distribution very popular in QUS techniques [19–26]. A review of the major distributions used for modeling signals scattered in soft tissue was presented by Destrempes and Cloutier [27].

Microstructural changes in the medium may also be captured and described by calculating entropy for ultrasound echo envelope value [28]. Information entropy has been defined by Shannon, and can be interpreted as a measure of the uncertainty of a given event [29]. In the analysis of signals scattered in a medium, entropy depends on the number of scatterers and the variance of their scattering properties [30].

In studies performed so far, tumors were assessed using classifiers based on parameters estimated from signals scattered on structures lying within the area identified as a lesion. However, both malignant and benign breast lesions are very heterogeneous, and differ in terms of the morphological features of the cells and, essentially, the stromal component proliferation. Based on ultrasound features, the most common subtype of breast cancer, LA (luminal A), is characterized by echogenic halo and post-acoustic shadowing [31]. This appearance is probably caused by the desmoplastic reactions in the tumor and surrounding tissue, and the invasion of malignant cells into fat tissue and elastic fibers. Another breast cancer subtype, LB (luminal B), reveals the absence of an echogenic halo, while the images of the HER2-amplified subtype show post-acoustic enhancement [31]. It was also proved that in shear wave elastography, the presence of the ‘stiff rim’ sign is a feature characteristic for malignant lesions [32]. On the other hand, fibroadenomas, especially large ones, could cause hyaline degeneration and presents ‘stiff rim’, but they do not infiltrate the surrounding tissue. It should be expected that data from the tissue surrounding the tumor, beside those from within the tumor, also provide diagnostically valuable information about the lesion. This is reflected, for instance, in the guidelines for BI-RADS classification, which refer to factors such as the posterior features of the image or calcification outside of the mass [5].

The purpose of this paper is to demonstrate that tissue surrounding the tumor is a source of ultrasound data useful for distinguishing malignant and benign breast lesions, which is especially relevant in the case of small tumors, as their size significantly restricts the amount of data that may be used for estimating tissue parameters. This data set limitation increases the estimation error for these parameters and reduces classification accuracy. The paper compares the results of tumor classification as malignant or benign using two parameters: weighted entropy and Nakagami distribution shape parameter. The choice of these parameters was dictated by the fact that they are used to classify neoplastic lesions. It is worth mentioning here that the purpose of our research was not to find new parameters that effectively differentiate benign and malignant changes. We wanted to show that the tissue surrounding the area of tumor indicated by the radiologist contains information about the malignancy of the tumor and can be used in the classification as well as the information from the lesion itself. Data for the estimation of weighted entropy and Nakagami distribution shape parameter were collected from the tumor area and from the surrounding tissue. Receiver operating characteristic (ROC) curves and areas under these curves (AUC) were used to evaluate the classification accuracy.

2 Materials and Methods

2.1 Data Collection

Ultrasound signals were registered at the Department of Radiology of Maria Skłodowska-Curie Memorial Institute of Oncology in Warsaw. The Institutional Review Board approved the study protocol. In this prospective study, all patients

signed the informed consent for breast US examination and for “backscatter US” statistical studies. Lesions of solid BI-RADS category 3, 4, or 5 were included in this study. The classical breast US examinations were performed according to the American College of Radiology BI-RADS guidelines and the Polish Ultrasound Society standards, using longitudinal and transverse scan planes [5, 33].

Radio frequency (RF) data were collected using the Ultrasonix SonicTOUCH[®] machine with an $L14 - 5/38$ linear probe. The transmit frequency was set at 10 MHz, and the captured signals were sampled at 40 MHz. The number of lines per RF image was 510. The classical transmit-receive scheme was used, with the focus set at the area within the tumor. Ultrasound scans of 130 tumors were collected, including 65 malignant and 65 benign cases. Each tumor was scanned in two orthogonal planes and each scan was treated as a separate measurement. Therefore, 260 RF data sets were used in calculations, 130 per category (malignant and benign). A biopsy was performed for each lesion and data were categorized on the basis of histopathology results. Tumor borders in the ultrasound images were delineated by an experienced physician.

Additionally, a data set excluding all tumors exceeding 7 mm in the vertical or horizontal dimension was analyzed separately. This size-restricted data set comprised 27 and 29 RF data sets of benign and malignant tumors, respectively.

2.2 ROI Selection

Classifiers were tested within the tumor area and in the tissue directly surrounding it. The surrounding area was delineated using a tumor mask, or an array with the same size as the RF image, containing ‘1’ values at locations corresponding to the lesion area, and ‘0’ values at the remaining locations. The mask was enlarged and then the original mask was subtracted from the enlarged one, producing the mask of the area surrounding the tumor (the ‘ring’). The width of the ring was ~ 6 mm. It was selected so as to optimize tumor classification accuracy. Classification accuracy was evaluated based on the area under the ROC curve (AUC), discussed in more detail in Sect. 2.4.

Five regions of interest (ROIs) for determining the averaged parameters were considered: the tumor area, the tumor and ring area, the ring area, the ring area above the midline of the tumor (‘upper half-ring’), and the ring area below the midline (‘lower half-ring’). An example B-mode image of a tumor with the ROIs marked is shown in Fig. 1.

2.3 Tissue Parameters

To characterize the tissue, the following parameters of the backscattered ultrasound signal envelope were analyzed: weighted entropy and the Nakagami distribution m -shape parameter. All calculations used in parameter estimation and classification were performed using commercial calculation software (MATLAB[®] 2015a). All parameters used for classification were estimated as follows. A parametric image was created, representing the spatial distribution of the parameter in the tumor area and its surroundings. Each point in the parametric image was

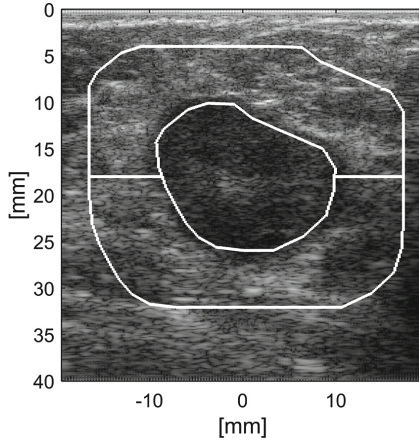


Fig. 1. Tumor an ROI borders marked with white lines.

assigned a value of the estimated parameter determined from a square window with a side length of 1 mm and a center at the location of the pixel. The relatively small window size made it possible to disregard the changes in signal amplitude resulting from variations of attenuation within the tumor. The window size used was in line with the window size for Nakagami distribution shape parameter estimation suggested in literature [34]. The window was moved in axial and lateral direction over the ROI with one sampling point step. Figure 2 shows parametric images of the tumor and its surroundings from Fig. 1 above. Subsequently, the average value of the parameter was determined for the entire ROI. These values were used for classification. The parameters and estimation procedures are described below.

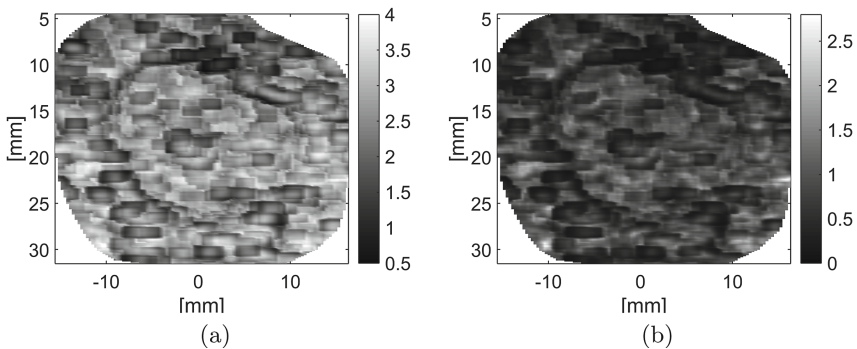


Fig. 2. Parametric images of (a) entropy and (b) m-shape parameter of Nakagami distribution calculated in the tumor and its surroundings.

Entropy. Shannon entropy is a measure of uncertainty of a random variable [29, 35]. Hughes proposed the use of entropy in quantitative assessment of changes in scattering medium structure [36]. Tsui studied the sensitivity of entropy and weighted entropy in detecting scatterer concentration, finding that while both parameters allow for assessing the spatial density of scatterers, weighted entropy is more sensitive to variation in scatterer concentration [28]. Tsui et al. also investigated the use of entropy for tissue characterization and breast tumor classification [30]. In the present study, weighted entropy [37] was used as a measure of local signal envelope heterogeneity. Weighted entropy H was estimated using formula (1):

$$H(A) = - \sum_{i=1}^n w(A_i) P(A_i) \log_2 P(A_i) \quad (1)$$

where n represented number of samples in a window, and w and P were the weight and the probability of amplitude value A , respectively. Normalized amplitudes A , were used as weights, which means that samples representing the highest amplitude in the analyzed sample set had a weight of 1. The probability P was estimated using the histogram method.

Nakagami distribution m -shape parameter. The Nakagami distribution is used for modeling the distribution of amplitudes of signals scattered in tissue, and can be used for the characterization of tissues, including breast tumor tissues [19]. The probability density function (PDF) for this distribution is given by formula (2) [18]:

$$P(A) = \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m A^{2m-1} \exp\left(-\frac{m}{\Omega}A^2\right) \quad (2)$$

where $\Gamma()$ is the gamma function, and m and Ω are the shape parameter and the scaling parameter associated with average signal power, respectively. In this paper, the Nakagami distribution m -shape parameter was used to characterize the statistical distribution of the signal envelope. The method of moments was used in the estimation [18], and the m parameter was estimated using formula (3):

$$m = \frac{\langle A^2 \rangle^2}{\sigma^2(A^2)} \quad (3)$$

where $\langle A^2 \rangle$ and $\sigma^2(A^2)$ are the mean and variance of amplitude squared, respectively.

The m parameter provides information on the signal envelope statistic. If $0 < m < 0.5$, the distribution is a Nakagami-Gamma distribution, corresponding to a low number of scatterers in the resolution cell. An m value between 0.5 and 1 corresponds to a generalized Rician distribution, and a random distribution and random cross sections of scatterers [19] or a periodical distribution of some scatterers at intervals of a quarter of the scattered pulse wavelength or denser

[38]. If m equals to 1, amplitudes have a Rayleigh distribution, corresponding to the scattering by a large number of point scatterers randomly distributed in the resolution cell. An m value above 1 corresponds to the Rician distribution, with both a large number of randomly distributed scatterers and scatterers distributed periodically at intervals equal to a multiple of half the scattered impulse wavelength. A description and analysis of the Nakagami distribution as a model for the amplitudes distribution of signal scattered in soft tissue was presented by Shankar [38].

2.4 Classifiers Evaluation

Classifiers were evaluated based on their receiver operating characteristics (ROCs) [39], and to be more precise, on the area under ROC curves (AUC). The ROC curve is a graphic representation of the performance of a binary classifier used for classifying cases into one of two classes – here, benign or malignant tumors – depending on the set discrimination threshold. The X axis shows the ratio of benign cases wrongly classified as malignant to all benign cases, i.e. the false positive rate or 1-specificity. By analogy, the Y axis shows the ratio of cases accurately classified as malignant to all malignant cases, i.e. the true positive rate or sensitivity. Any arbitrary discrimination threshold is associated with a certain number of true and false positives, marked by a point in the chart. The ROC curve is plotted by changing the discrimination threshold. The area under the curve (AUC) is a numerical parameter showing how well a classifier performs, and is often used in classifier evaluation. It can range between 0 and 1. Values below 0.5 indicate that the classifier performs worse than random guessing. Thus, it can be assumed that for realistic classifiers, the AUC should be between 0.5 (for random classification) and 1 (for an ideal classifier) [39].

3 Results

Areas under ROC curve (AUC values) together with estimation errors, for the studied parameters estimated for all the analyzed ROIs are shown in Table 1. Comparisons between ROC curves for parameters estimated within the tumor and in the ring outside of the tumor are shown in Fig. 3.

Table 1. Values of areas under the ROC curves (AUC)

	Tumor	Tumor & ring	Ring	Upper half-ring	Lower half-ring
Entropy	0.67(3)	0.78(3)	0.78(3)	0.74(3)	0.79(3)
Nakagami	0.68(3)	0.79(3)	0.79(3)	0.76(3)	0.80(3)

AUC values estimated in the set of data from tumors not exceeding 7 mm in the vertical or horizontal dimension are listed in Table 2. Comparisons between ROC curves for parameters estimated within the tumor and in the ring outside of the tumor for these cases are shown in Fig. 4.

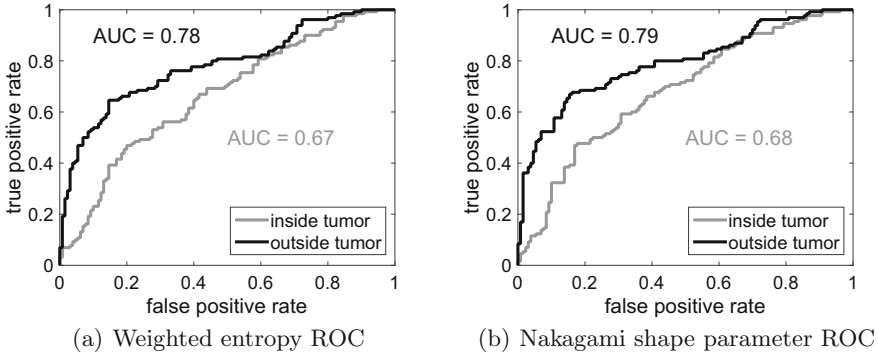


Fig. 3. Comparison of ROC curves of (a) weighed entropy and (b) shape parameter.

Table 2. Values of areas under the ROC curves (AUC) estimated for size restricted tumors

	Tumor	Tumor & ring	Ring	Upper half-ring	Lower half-ring
Entropy	0.53(7)	0.72(7)	0.72(7)	0.68(7)	0.73(7)
Nakagami	0.54(7)	0.73(7)	0.73(7)	0.72(7)	0.72(7)

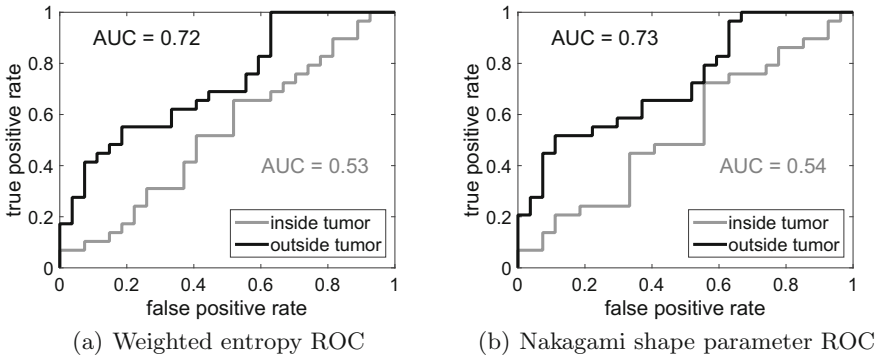


Fig. 4. Comparison of ROC curves of (a) weighed entropy and (b) shape parameter estimated for size restricted tumors.

4 Discussion

As shown in Table 1 AUC values for weighted entropy and Nakagami distribution shape parameter were similar. Significant differences were found between AUC values for parameters estimated in various ROIs. It is worth emphasizing that the most common approach, i.e. the use of the tumor area only, led to the worst results of tumor classification (AUC equal to 0.67(3) and 0.68(3) for entropy and Nakagami m parameter respectively). All the considered ROI cases that included surrounding areas resulted in increased AUC values. The best results

were obtained for the lower half-ring only (AUC increases by $\sim 18\%$), but any ROI configuration containing tissue surrounding the tumor led to a comparable improvement. The enhancement of classification by parameter estimation in the ring surrounding the tumor is illustrated by the ROC curves (Fig. 3).

AUC values for the data set only comprising small tumors, not exceeding 7 mm in the front-to-back (vertical) or side-to-side (horizontal) dimension, are shown in Table 2. There is a tendency similar to that found in the full data set (Table 1), i.e. estimation of parameters in the area containing the tissue surrounding the tumor is more accurate than in the tumor area only. Notably, for small tumors, estimations made in the tumor area only were associated with AUC values close to 0.5, which corresponds to random guessing.

5 Conclusions

The results reported here demonstrate that in the case of breast tumor classification using averaged statistical parameters – entropy and Nakagami distribution shape parameter – it was beneficial to extend the ROI beyond the lesion borders. This observation is consistent with the assumption that the tumor affects the surrounding tissue and its properties, which is reflected in the statistical parameters of the scattered ultrasound signal. Moreover, it is known that the majority of malignant tumors have indistinct, angular, microlobulated or spiculated margins or any combination of these features. They interact with the surrounding parenchymal tissue which induced the desmoplastic reaction, especially in low-grade breast cancers [40].

Classification based on parameters estimated within the tumor and the surrounding tissue, or even in the surrounding tissue only, was markedly more accurate than that based on data from within the tumor only. Another conclusion is that very precise delineation of the tumor boundaries may be unnecessary in the case of the studied parameters. This can be significant for cases where tumor boundaries are identified by a less experienced radiologist, as well as for the interpretation of ultrasound images using computer-aided diagnosis (CADx) systems.

The obtained results demonstrated that the extension of the ROI was particularly beneficial in the case of very small tumors, where parameter estimation error was large, as the estimates couldn't be appropriately averaged. In these cases, the larger area allows for reducing the estimation error and thus enhances classification. This is an important observation, as the advances in ultrasound techniques and the increasingly universal screening result in the detection of ever smaller tumors. Hence the importance of methods supporting the physician in the classification of such small focal lesions.

Acknowledgements. This study was supported by the National Science Centre, Poland, grants 2016/23/B/ST8/03391, 2016/21/N/ST7/03029 and 2014/13/B/ST7/01271. The project was implemented using the infrastructure of CePT, Operational Program “Innovative economy” for 2007–2013.

References

1. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F.: Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**(5), E359–E386 (2015)
2. Wojciechowska, U., Olasek, P., Czauderna, K., Didkowska, J.: Cancer in Poland in 2014. Centrum Onkologii-Instytut im. Marii Skłodowskiej-Curie (2016)
3. Kolb, T.M., Lichy, J., Newhouse, J.H.: Comparison of the performance of screening mammography, physical examination, and breast us and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* **225**(1), 165–175 (2002)
4. Mandelson, M.T., Oestreicher, N., Porter, P.L., White, D., Finder, C.A., Taplin, S.H., White, E.: Breast density as a predictor of mammographic detection: comparison of interval-and screen-detected cancers. *J. Natl Cancer Inst.* **92**(13), 1081–1087 (2000)
5. Mendelson, E., Böhm-Vélez, M., Berg, W., Whitman, G., Feldman, M., Madjar, H., Rizzatto, G., Baker, J., Zuley, M., Stavros, A., Comstock, C., Van Duyn Wear, V.: ACR BI-RADS® ultrasound. ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System, vol. 149. American College of Radiology, Reston (2013)
6. D’Astous, F.T., Foster, F.S.: Frequency dependence of ultrasound attenuation and backscatter in breast tissue. *Ultrasound Med. Biol.* **12**(10), 795–808 (1986)
7. Nam, K., Zagzebski, J.A., Hall, T.J.: Quantitative assessment of in vivo breast masses using ultrasound attenuation and backscatter. *Ultrason. Imaging* **35**(2), 146–161 (2013)
8. Lizzi, F.L., Astor, M., Liu, T., Deng, C., Coleman, D.J., Silverman, R.H.: Ultrasonic spectrum analysis for tissue assays and therapy evaluation. *Int. J. Imaging Syst. Technol.* **8**(1), 3–10 (1997)
9. Moon, W.K., Lo, C.M., Chang, J.M., Huang, C.S., Chen, J.H., Chang, R.F.: Quantitative ultrasound analysis for classification of bi-rads category 3 breast masses. *J. Digit. Imaging* **26**(6), 1091–1098 (2013)
10. Tadayyon, H., Sadeghi-Naini, A., Czarnota, G.J.: Noninvasive characterization of locally advanced breast cancer using textural analysis of quantitative ultrasound parametric images. *Transl. Oncol.* **7**(6), 759–767 (2014)
11. Cai, L., Wang, X., Wang, Y., Guo, Y., Yu, J., Wang, Y.: Robust phase-based texture descriptor for classification of breast ultrasound images. *Biomed. Eng. Online* **14**(1), 26 (2015)
12. Cheng, H.D., Shan, J., Ju, W., Guo, Y., Zhang, L.: Automated breast cancer detection and classification using ultrasound images: a survey. *Pattern Recogn.* **43**(1), 299–317 (2010)
13. Dutt, V., Greenleaf, J.F.: Ultrasound echo envelope analysis using a homodyned K distribution signal model. *Ultrason. Imaging* **16**(4), 265–287 (1994)
14. Hruska, D.P., Oelze, M.L.: Improved parameter estimates based on the homodyned K distribution. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **56**(11), 2471–2481 (2009)
15. Hruska, D.P.: Improved techniques for statistical analysis of the envelope of backscattered ultrasound using the homodyned K distribution. Master’s thesis, University of Illinois at Urbana-Champaign (2009)

16. Trop, I., Destremes, F., El Khoury, M., Robidoux, A., Gaboury, L., Allard, L., Chayer, B., Cloutier, G.: The added value of statistical modeling of backscatter properties in the management of breast lesions at us. *Radiology* **275**(3), 666–674 (2014)
17. Byra, M., Nowicki, A., Wróblewska-Piotrkowska, H., Dobruch-Sobczak, K.: Classification of breast lesions using segmented quantitative ultrasound maps of homodyned K distribution parameters. *Med. Phys.* **43**(10), 5561–5569 (2016)
18. Nakagami, M.: The m-distribution—a general formula of intensity distribution of rapid fading. In: *Statistical Method of Radio Propagation* (1960)
19. Shankar, P.M., Dumane, V.A., Reid, J.M., Genis, V., Forsberg, F., Piccoli, C.W., Goldberg, B.B.: Classification of ultrasonic B-mode images of breast masses using nakagami distribution. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **48**(2), 569–580 (2001)
20. Gefen, S., Tretiak, O.J., Piccoli, C.W., Donohue, K.D., Petropulu, A.P., Shankar, P.M., Dumane, V.A., Huang, L., Kutay, M.A., Genis, V., et al.: ROC analysis of ultrasound tissue characterization classifiers for breast cancer diagnosis. *IEEE Trans. Med. Imaging* **22**(2), 170–177 (2003)
21. Tsui, P.H., Chang, C.C., Ho, M.C., Lee, Y.H., Chen, Y.S., Chang, C.C., Huang, N.E., Wu, Z.H., Chang, K.J.: Use of nakagami statistics and empirical mode decomposition for ultrasound tissue characterization by a nonfocused transducer. *Ultrasound Med. Biol.* **35**(12), 2055–2068 (2009)
22. Tsui, P.H., Yeh, C.K., Liao, Y.Y., Chang, C.C., Kuo, W.H., Chang, K.J., Chen, C.N.: Ultrasonic nakagami imaging: a strategy to visualize the scatterer properties of benign and malignant breast tumors. *Ultrasound Med. Biol.* **36**(2), 209–217 (2010)
23. Liao, Y.Y., Tsui, P.H., Li, C.H., Chang, K.J., Kuo, W.H., Chang, C.C., Yeh, C.K.: Classification of scattering media within benign and malignant breast tumors based on ultrasound texture-feature-based and nakagami-parameter images. *Med. Phys.* **38**(4), 2198–2207 (2011)
24. Ma, H.Y., Lin, Y.H., Wang, C.Y., Chen, C.N., Ho, M.C., Tsui, P.H.: Ultrasound window-modulated compounding nakagami imaging: resolution improvement and computational acceleration for liver characterization. *Ultrasonics* **70**, 18–28 (2016)
25. Tsui, P.H., Wan, Y.L.: Application of ultrasound nakagami imaging for the diagnosis of fatty liver. *J. Med. Ultrasound* **24**(2), 47–49 (2016)
26. Dobruch-Sobczak, K., Piotrkowska-Wróblewska, H., Roszkowska-Purska, K., Nowicki, A., Jakubowski, W.: Usefulness of combined bi-rads analysis and nakagami statistics of ultrasound echoes in the diagnosis of breast lesions. *Clin. Radiol.* **72**(4), 339–e7 (2017)
27. Destremes, F., Cloutier, G.: A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope. *Ultrasound Med. Biol.* **36**(7), 1037–1051 (2010)
28. Tsui, P.H.: Ultrasound detection of scatterer concentration by weighted entropy. *Entropy* **17**(10), 6598–6616 (2015)
29. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423, 623–656 (1948)
30. Tsui, P.H., Chen, C.K., Kuo, W.H., Chang, K.J., Fang, J., Ma, H.Y., Chou, D.: Small-window parametric imaging based on information entropy for ultrasound tissue characterization. *Sci. Rep.* **7**, 41004 (2017)
31. Zhang, L., Li, J., Xiao, Y., Cui, H., Du, G., Wang, Y., Li, Z., Wu, T., Li, X., Tian, J.: Identifying ultrasound and clinical features of breast cancer molecular subtypes by ensemble decision. *Sci. Rep.* **5**, 11085 (2015)

32. Zhou, J., Zhan, W., Chang, C., Zhang, X., Jia, Y., Dong, Y., Zhou, C., Sun, J., Grant, E.G.: Breast lesions: evaluation with shear wave elastography, with special emphasis on the “stiff rim” sign. *Radiology* **272**(1), 63–72 (2014)
33. Jakubowski, W., Dobruch-Sobczak, K., Migda, B.: Standards of the polish ultrasound society-update. Sonomammography examination. *J. Ultrason.* **12**(50), 245 (2012)
34. Tsui, P.H., Ma, H.Y., Zhou, Z., Ho, M.C., Lee, Y.H.: Window-modulated compounding nakagami imaging for ultrasound tissue characterization. *Ultrasonics* **54**(6), 1448–1459 (2014)
35. Wang, Q.A.: Probability distribution and entropy as a measure of uncertainty. *J. Phys. A: Math. Theor.* **41**(6), 065004 (2008)
36. Hughes, M.S.: Analysis of digitized waveforms using shannon entropy. *J. Acoust. Soc. Am.* **93**(2), 892–906 (1993)
37. Guiaşu, S.: Weighted entropy. *Rep. Math. Phys.* **2**(3), 165–179 (1971)
38. Shankar, P.M.: A general statistical model for ultrasonic backscattering from tissues. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **47**(3), 727–736 (2000)
39. Fawcett, T.: An introduction to roc analysis. *Pattern Recogn. Lett.* **27**, 861–874 (2006)
40. Stavros, A.T.: *Breast Ultrasound*. Lippincott Williams & Wilkins, Philadelphia (2004)



A Texture Analysis Approach for Spine Metastasis Classification in T1 and T2 MRI

Mohamed Amine Larhman¹✉, Saïd Mahmoudi¹, Stylianos Drisis²,
and Mohammed Benjelloun¹

¹ Computer Science Department, Faculty of Engineering, University of Mons,
Mons, Belgium

mohamedamine.larhman@umons.ac.be

² Radiology Department, Jules Bordet Institute, Université Libre de Bruxelles,
Brussels, Belgium

Abstract. This paper presents a learning based approach for the classification of pathological vertebrae. The proposed method is applied to spine metastasis, a malignant tumor that develops inside bones and requires a rapid diagnosis for an effective treatment monitoring. We used multiple texture analysis techniques to extract useful features from two co-registered MR images sequences (T1, T2). These MRIs are part of a diagnostic protocol for vertebral metastases follow up. We adopted a slice by slice MRI analysis of 153 vertebra region of interest. Our method achieved a classification accuracy of $90.17\% \pm 5.49$, using only a subset of 67 relevant selected features from the initial 142.

Keywords: MRI · Machine learning · Texture analysis · GLCM · LBP

1 Introduction

Bone metastases represents a frequent complication in case of systemic cancer [1]. They are malignant tumors that develop inside bones and results from the migration of cancer cells via blood or lymph circulation into bone nucleus. Human spine is considered as one of the most frequent site of bone metastasis [2]. Statistics show that 80% of bone metastases originating primary tumor came from breast or prostate cancers [3]. Spine metastases can take different forms: osteolysis, osteocondensation or a mixed form of both [4].

Patients with spine metastasis are subject to continuous pain, risk of fractures and side effects caused by possible treatments with chemo-therapeutic agents [5]. This kind of high cost treatment is toxic and still unclear. Thus, rapid evaluation of its efficiency is an important step for clinical decision making. Clinical diagnosis of spine metastasis can be made with different exams such as MRI, CT-scan and PET.

2 Related Works

In the literature, we found different studies related to bone metastases analysis using a computer aided diagnostic (CAD) system. In this context, Jerebko et al. [6] proposed a prototype CAD system for vertebra metastases detection in MRI. They used a geometric model to detect the spine curve, which is used to extract regions of interest. The detection of metastases is performed by the extraction of different image features and a classification by Discriminant Linear Analysis [7,8]. Burns et al. [9] proposed a system for quantitative detection and characterization of osteoblast metastases from thoracic and lumbar vertebrae in CT-scan. Detection of metastases is performed by using an SVM classifier that takes 25 features (morphologic, attenuation and positions). The reported detection rate was 90%. Huang and Chian [10] and Yao et al. [11] used similar approaches for the detection of lytic metastases in CT-scan.

Another approach based on Deep Learning [12] techniques presented by Roth et al. [13] showed the efficiency of convolution neural networks (CNNs) for the identification of lytic vertebral metastases in CT-scan. In this kind of approach, useful image features are automatically learned from the image data base. Despite the high efficiency of deep learning and its breakthrough in multiple computer vision applications, this approach needs large data sets for learning a good feature representation.

Most of the studies related to metastases analysis are mainly concerned with CT-Scan images. However, this medical imaging technique represents a potential risk for human health due to radiation exposure. For example, during a lung CT-Scan exam a patient receives a dose of X-rays radiation 10 to 100 times greater than a pulmonary radio. Therefore, in some cases, MRI exam can be a major alternative to CT-scan. MR images are characterized by a large variation in gray level compared to other medical imaging modalities. This variation represents the different soft tissues in human body, which appear as image textures. In medical image analysis field, texture analysis has been used, as a new bio-marker, for the characterization of tumors and lesions.

Recent techniques, based on texture analysis, have been applied for liver and kidney metastases diagnosis on CT-scans. Haider et al. [14] evaluated the potential of texture analysis for predicting the response to metastatic kidney cell therapy. They concluded that this approach improves the results of patient survival prediction. Rao et al. [15] compared the effectiveness of texture analysis to size measurements by RECIST criteria for liver metastases. They concluded that texture features (entropy and uniformity) outperform size measurements for the evaluation of the response to chemotherapy in CT-scans. Texture based analysis enabled to extract new clinical biomarkers other than the classical tumor size.

In this context, Vallieres et al. [16] used multiple texture features extracted from both MRI and FDG-PET for the early evaluation of pulmonary metastases in soft tissue sarcomas (STS). This work, showed the effectiveness of applying texture analysis as bio-markers for the classification of aggressive tumors using a learning based scheme. Holli et al. [17] applied texture analysis for cranial

trauma in MR images. They concluded that texture analysis can be applied as a **new early diagnostic tool** for brain lesions follow up. Similarly, authors in [18] evaluated textures analysis for the characterization of different types of breast cancer in MRI.

In case of spine metastases follow up, the effectiveness of medical treatments on the evolution of metastases is difficult to analyze which requires a CAD system to evaluate the response to treatment. In this context, MRI has the advantage of being a non-invasive imaging technology (no radiation) for human health, and it allows a better differentiation between healthy and pathological tissues. In addition, it is the most sensitive exam to evaluate the early response to chemotherapy. For these reasons, MRI is highly recommended in oncology clinic for metastases treatment monitoring and follow up [19,20]. However, few works have been reported in the literature for the development of a CAD dedicated to the analysis of bone metastases in MR images. The latter remains an open subject of research.

This work is concerned with the development of a Computer-aided diagnosis system for the analysis and follow up of vertebra metastases in T1 and T2 MRI. We propose a learning approach to evaluate the effectiveness of texture analysis to classify abnormal vertebrae. We used multiple image texture features, extracted from image histogram, gray-level cooccurrence matrix (GLCM) and local binary patterns (LBP). Using a low size data base for a machine learning problem remains a challenging task which have been tackled in our case using a grid search nested cross validation to setup an SVM classifier.

3 Materials and Methods

The proposed approach is based on the following steps (*cf.* Fig. 1):

- Data preprocessing: including ground truth preparation, image co-registration and ROIs extraction
- Feature extraction: multiple texture analysis techniques (histogram, GLCM and LBP) calculated from ROIs
- Classifier learning: a machine learning pipeline including data normalization, cross validation and feature selection.

3.1 Data Preprocessing

The MRI case used in this work were obtained from Jules Bordet Institute¹, it contains two types of MRI sequences; T1 and T2 spine sagittal sections of a patient with multiple metastases. These MRIs are part of a diagnostic protocol for vertebral metastases follow up. They are characterized by a small thickness in the coronal direction centered around vertebrae. The database has been

¹ Multidisciplinary and autonomous hospital totally dedicated to cancer located in Brussels, Belgium.

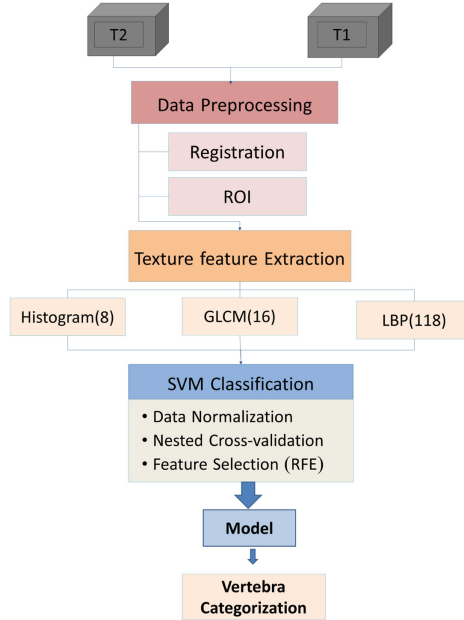


Fig. 1. Overview of the pipeline of classification model selection: pre-treatment of MR images (T1 and T2), then the extraction of texture features, and finally the learning step using SVM method

annotated and validated with the help of an experimented radiologist. Therefore, in T1-MRI each vertebra body slice has been annotated and labeled as pathological or healthy. We used T1-MRI since vertebra shape and edges are better highlighted in this sequence compared to T2-MRI. An example of these annotations is shown in Fig. 2.

Then, we conducted an affine registration of T1 and T2 sequences. We started with a manual selection of the slices that show entire vertebrae from both sequences. This step enabled to align the two volumes in the coronal direction not considered by registration method. Then, an affine registration was applied to the selected slices. T1 being the fixed volume, the deformation was applied to the T2 volume. Figure 2 shows a registration example of two central slices from T1 and T2 MRI.

The adopted registration technique is based on similarity measurement which quantify a distance metric between the reference image and transformed image. In this work, we used an affine transformation based on the standardized mutual information similarity measure [21]. The latter is commonly used for the registration of images of multi-modalities or multi-sequences [22]. We used the module proposed by Johnson et al. [23] (3D BRAINSFit) available in 3DSlicer² software.

² www.slicer.org.

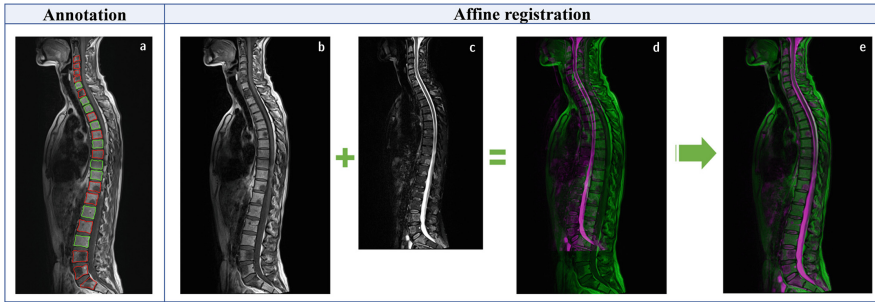


Fig. 2. (a) Example of vertebra annotations of a central slice from T1-MRI: healthy vertebra in green and pathological vertebrae in red. Affine Registration on two MR central slices: (b) T1 and (c) T2, of a patient with multiple metastases. (d) Shows an overlay of the input images. (e) Shows the final result, T1 in green and T2 (transformed) in purple (Color figure online)

Then, the annotation masks have been used to extract local regions of interest from registered images. These ROIs represent both healthy and pathological vertebra volumes. Each sagittal slice is used as a separate ROI for a slice by slice analysis approach. The data were labeled into two classes: Pathologic (87 ROIs) and Healthy (66 ROIs), a total of 153 ROIs.

3.2 Texture Feature Extraction

Image Texture is defined as spatial variation of pixel intensities [24], it represents the systematic local variation of image values [25]. Several image texture analysis approaches have been used in the literature including statistical approaches (histogram and GLCM [26]), model-based approaches (fractal models and Markov fields) and transformation based approaches (Fourier transform, Wavelet transform or Gabor filtering).

Statistical texture analysis has been widely used for medical image analysis and in particular for cancer monitoring, for example, to evaluate the response to radiotherapy [24, 27]. This approach uses statistical parameters to characterize the structure of an image or a region of interest.

First-order statistics, based on histogram, provide an overall evaluation of the distribution of pixel intensities in an image. In our experimentation, we used 4 first-order statistic measures based on gray level histogram (Variance, Skewness, Kurtosis, Mean). Second order statistics evaluate the spatial relationship between pixels, they are extracted from grey-level cooccurrence matrix (GLCM) [26]. The latter “computes the frequency of which pairs of pixels with specific intensity values occurs in an image with a given spatial relationship”. We extracted 8 texture features from GLCM (Energy, Contrast, Correlation, Homogeneity, Variance, Sum Average, Entropy, Dissimilarity) as proposed in [16].

Local Binary Patterns (LBP) [28] is an operator which characterizes the local structure of image texture. This method calculates the binary difference between the gray level of a given pixel and his neighbors. LBP is known for its robustness to the monotonic variations of illumination, which makes it suitable for MRI data, and for its simple computational complexity. We found many applications of LBP method to facial recognition [29], and some recent applications in the field of medical image analysis for brain tumors recognition [30] and analysis [31]. In the context of our experimentation, 59 uniform LBP codes [32] have been used. Table 1 shows the total number of extracted texture features on each MRI sequence (T1 and T2).

Table 1. Numbers of extracted texture features by type on each MRI sequence

Texture methods	Type	Number			Total
		T1	T2	T1 + T2	
First-order	Global	4	4	8	142
Second-order	GLCM	8	8	16	
LBP	Uniform	59	59	118	

Details of the extraction of second-order texture features are as follow; the size of co-occurrence matrix is $N_g = 8$. In order, to achieve a rotation-invariant feature extraction, we computed four GLCM matrices according to the four angles: $0, \pi/4, \pi/2, 3\pi/4$. Texture features are extracted from each matrix and then averaged.

We configured the LBP function as follow; number of neighborhood $p = 8$ and radius $r = 1$. We used uniform models, this choice reduces the number of LBP codes from $2^8 = 256$ to 59. Therefore, the size of the extracted LBP histogram from each region of interest is $K = 59$.

Figure 3 shows the normalized LBP histogram of the entire data set for the two classes. We notice a high value of the LBP codes which are in the two borders of both histograms. These codes represent the homogeneous regions, their values are different from one class to another. This figure shows only the global histograms of all regions of interest. Furthermore, individual LBP histograms of each region of interest will be used for vertebra categorization step.

3.3 Classifier Learning

In this section, we describe our supervised learning approach based on multiple SVMs [33] classifiers. SVMs method is often used for medical image classification, one advantage of this technique is its capacity of generalization even with small learning sets [34]. Also, it is efficient for the analysis of data sets with large spaces (large number of features). As a first step, we performed a preliminary test on several existing methods: SVMs, LDA [35], k-nearest neighbors (KNN) [36], RF [37], SVMs scores surpassed the other methods in our case.

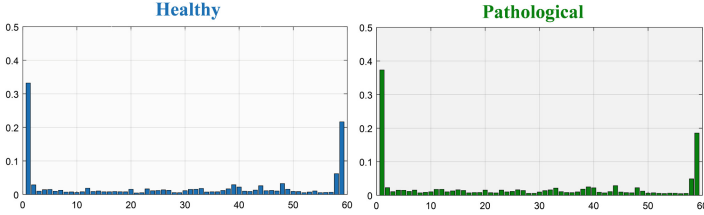


Fig. 3. Normalized mean histogram of LBP codes for the two learning classes, healthy and pathological, in T1 images

For the evaluation of the proposed method, we conducted a grid search comparison of the performance of different SVM kernels: linear, Gaussian and polynomial.

Data standardization: SVMs, like other supervised learning methods, is sensitive to, and it performs well on standardized data. Standardization enables to limit the variance between features, prevents the dominance of large values, and minimize numerical complexity during the calculation [38]. This type of data transformation, unlike normalization technique, enables to keep useful information on outliers while making the algorithm less sensitive to them. The transformation consists of subtracting the mean and dividing by the standard deviation of each feature column, as shown in the following equation:

$$standardization(x) = \frac{x - \mu}{\sigma} \quad (1)$$

where μ is the mean of feature column, and σ its standard deviation.

μ and σ are calculated on the training data and the same values are used for the transformation of test data.

Cross-validation: In machine learning pipeline, validation of a model on new data which have not been used for training, is an important step for evaluating its ability for generalization. For this purpose, the initial data can be divided into 2 subgroups; learning set and test set. Furthermore, tuning the input parameters have also a considerable impact on the final performance on test data [39]. One solution is to subdivide the initial data into three subgroups: learning, validation, and testing. In this way, the models are first trained on learning set, then the parameter tuning is done on validation set, and finally the performance is calculated on test set to prevent overfitting.

To make this process robust and less sensitive to data subdivision, *k-fold* cross-validation can be performed. In our case, we used a more generalized scheme based on nested cross validation (NCV) method. This method enables to calculate a better model performance on tested data [40]. Indeed, NCV varies the initial subdivision of test and validation data each time. This is done by an external loop which splits the data into learning folds and a test fold and then an internal loop, on learning folds, used for parameter tuning [39]. We used a

case of NCV called 5×2 cross validation, where the outer loop splits the data into five folds and the inner loop into two folds.

In addition, we used a stratified cross validation technique which keeps the same percentage of the two classes in each fold. As a result, each fold is representative of the structure of the database. Kohavi [41] have shown that the use of stratified cross validation can guarantee a better bias-variance trade-off, where the bias represents the model error and the variance is related to sensitivity to subdivision of the data.

Model selection: SVM method requires several input parameters, namely: regularization constant C , kernel type, kernel coefficient γ and *Degree*. The choice of these parameters has an impact on the final classifier score.

Optimizing the choice of these parameters can be represented as a search problem. To solve it, there are two known strategies; search grid and random search. Search grid is an approach that evaluates all the combinations from a grid of parameters. Unlike random search which uses a random distribution of parameter. Search grid is suitable in our case since it allows to specify significant values for a reasonable calculation time. We used a logarithmic scale to set a grid search as recommended in [38]. Table 2 shows the used grid search. Note that the *Degree* parameter is used only by the polynomial kernel, and the kernel parameter γ is not used by the linear model. The global accuracy *ACC* is often used as a metric to evaluate a classification model, its formula is written as follows:

$$ACC = \frac{TP + TN}{Total} = 1 - ERR \quad (2)$$

where TP are true positives (fraction of positives actually detected) and TN are true negatives (fraction of negatives that are correctly detected).

Table 2. The used grid search for parameter tuning of SVM classifiers

Parameters	Values
<i>Kernel</i>	Linear, polynomial, Gaussian
<i>Degree</i>	{2, 3, 4}
C	$\{10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^2, 10^3\}$
γ	$\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$

4 Experimental Results

In this section, we present the experimental results related to the construction and the evaluation of vertebra categorization models from T1 and T2 MRI. Vertebra categorization can be represented as a classification problem of healthy and pathological vertebrae (with bone metastases). This classification is applied to multiple image features extracted from vertebra ROIs.

Experimentation have been conducted using image data from both T1 and T2 MRI. The size of each feature space is shown in Table 1. We compare the performance of the proposed method on each data set (T1, T2 and T1 + T2) using first-order, GLCM and LBP image features.

Table 3 summarizes the results of the evaluation of classification models using NCV method based on global accuracy score. We can notice that in T1 data set, training a model on the combination of the three image features enables to improve the individual scores of the models trained on each one separately, going from 76.42% and 70.03% to 82.88%. Likewise, in T2 data set, the model trained on the three image features performs well with less standard deviation $88.86\% \pm 3.35$.

The model with the highest score is the one trained on both T1 and T2 data set using the three images features with an accuracy of $90.08\% \pm 5.65$. We notice that adding LBP features enables to jump from 86.77% to 90.08% even if these features does not perform well individually.

Table 3. Results of the evaluation, with accuracy score (\pm standard deviation), of the proposed method on both T1 and T2 MRI, using histogram, GLCM and LBP based image features

		ACC	\pm Std
T1	Global + GLCM	76.42%	± 8.30
	LBP	70.03%	± 4.98
	Total	82.88%	± 5.57
T2	Global + GLCM	88.79%	± 5.84
	LBP	71.05%	± 10.34
	Total	88.86%	± 3.35
T1 + T2	Global + GLCM	86.77%	± 7.10
	LBP	79.79%	± 5.26
	Total	90.08%	± 5.65

Feature selection: We showed that the combination of first-order, GLCM and LBP image features in T1 and T2 MRI improved the classification scores. However, increasing the dimension of the problem (142 features) implies both higher computation cost and risk of overfitting. In Fig. 4, we study the variation of the mean computation time of a single SVM prediction as a function of the number of used features. We notice that computation time is influenced by two factors: the number of features and the complexity of SVM kernel. For example, for the case of 142 features, a linear kernel is about 5 times faster than the polynomial and Gaussian kernels.

One solution is to apply a feature selection method which consist of selecting a subset the most relevant features for our prediction problem. *Recursive Feature Elimination* (RFE or SVM-RFE) [42] is a feature selection technique that is

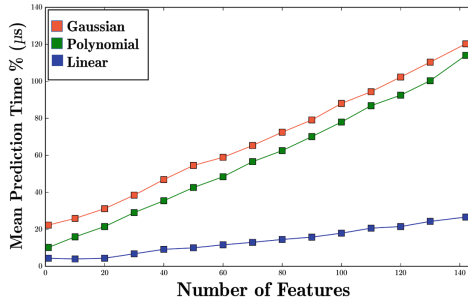


Fig. 4. Influence of SVM kernel type and number of features on the mean computation time of prediction

often combined with SVM classifier. RFE train iteratively a linear SVM model and removes each time the least relevant features based on their coefficients (calculated with SVM). This technique has the advantage of making selection at the same time as model learning.

We applied SVM-RFE on T1 + T2 including all images features (142). Table 4 shows the impact of adding feature selection on model accuracy. We notice a slight gain in accuracy and a decrease in the number of selected features, ranging from 143 to 67. Furthermore, Fig. 5 shows the evolution of the accuracy of SVM model with linear kernel (during SVM-RFE process) as a function of the number of selected features. We notice that the accuracy reaches its maximum of 87.48% with 67 features and decreases slightly thereafter. The classification model becomes stable after 67 features. The 67 selected features are then used to retrain multiple SVM models on the grid search with NCV, the final accuracy is $90.17\% \pm 5.49$.

Table 5 shows the distribution of selected features using RFE-SVM method. We notice that the selected features came from both T1 and T2 datasets and are well distributed on the three images features extraction techniques used in this experimentation.

4.1 Performance of Selected Models

In order to evaluate and compare the performance of the presented classification models, we use ROC curve (Receiver Operating Characteristic). ROC curve is a tool that allows to analyze the behavior of a classifier model with different thresholds (values for which an instance belongs to a given class). It is in the form of a curve that gives the rate of true positive as a function of the false positive rate. ROC curve is also used to calculate the AUC (Area Under Curve). A good classifier is characterized by a larger AUC.

Figure 6 compares the performance, based on ROC curve and AUC, of the classification models trained on T1 and T2 datasets individually and combined including all images features and selected features with RFE-SVM. We can notice that the model trained on T1 + T2 with RFE-SVM obtains the best AUC followed by the model trained on T1 + T2 including all features.

Table 4. Feature selection results: global accuracy before and after applying SVM-RFE with standard deviation and the final number of selected features

	Acc	Std	Features Nbr
T1 + T2	90.08%	±5.65	143
RFE T1 + T2	90.17%	±5.49	67
Gain	+0.90%	-	-76

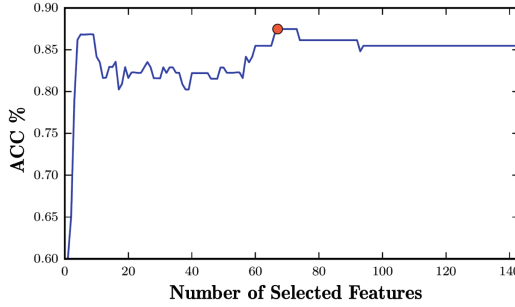


Fig. 5. The curve of linear SVM accuracy of the model trained on T1 + T2 as a function of the number of selected features using RFE-SVM method. The highest accuracy (during SVM-RFE process) of 87.48% is observed with 67 features (point in red) (Color figure online)

Table 5. Distribution of selected features using RFE-SVM method

Features type	T1		T2	
	GLCM (12)	LBP (59)	GLCM (12)	LBP (59)
Nbr of selected features	3	27	7	30

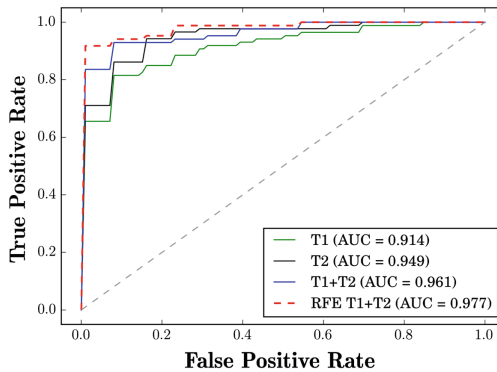


Fig. 6. Comparison ROC curves for SVM models trained on T1, T2, T1 + T2 with and without feature selection

5 Discussion and Conclusion

In this paper, we presented a method for the categorization of healthy and pathological vertebrae from both T1 and T2 MRI sequences. We formulated this task as a classification approach based on multiple texture features. We obtained the highest accuracy of 90.08% on T1 + T2 image data set by using the whole set of features. We observed a slight improvement after applying feature selection, reaching an accuracy of 90.17% with only a subset of 67 features.

In this work, we demonstrated how texture analysis, which have been proven to be effective for early diagnostic from medical images, can be applied for the categorization of vertebra, with spine metastases, on T1 and T2 MRI. Otherwise, we note the difficulty of medical data collection and annotation. As a result, we have been challenged with a low size data set. To overcome this limitation we adapted our solution; with the choice of SVM as a classification method (efficient in case of small data sets), a binary classification problem and finally using nested cross validation scheme (which maximizes the use of data for evaluation).

Vertebra categorization can be applied to identify pathological vertebra based on a first segmentation step, a task which have been widely reported in the literature [43]. A second use case is the medical follow-up of vertebrae metastases with a diffuse (non-homogeneous) morphology where it is necessary to analyze the whole vertebra instead of the volume of the lesion.

Finally, we note that the obtained results thus far were considered promising by practitioners. Nevertheless, a validation on a larger data set is necessary to allow a clinical feedback on the effectiveness of the method. Further analysis can be performed on a volume of interest instead of a 2D slide by slide analysis. Also, a more extensive comparison of features extraction techniques, including deep features, can be performed on larger data set.

References

1. Roodman, G.D.: Mechanisms of bone metastasis. *N. Engl. J. Med.* **350**(16), 1655–1664 (2004). PMID: 15084698
2. Guillevin, R., et al.: Spine metastasis imaging: review of the literature. *J. Neuro-radiol.* **34**(5), 311–321 (2007)
3. Coleman, R.: Metastatic bone disease: clinical features, pathophysiology and treatment strategies. *Cancer Treat. Rev.* **27**(3), 165–176 (2001)
4. Junsung Choi, M.R.: Diagnostic imaging and image-guided therapy of skeletal metastases. *Cancer Control* **19**(2), 102–112 (2012)
5. Gerszten, P.C., et al.: Radiosurgery for spinal metastases: clinical experience in 500 cases from a single institution. *Spine* **32**(2), 193–199 (2007)
6. Jerebko, A.K., et al.: Robust parametric modeling approach based on domain knowledge for computer aided detection of vertebrae column metastases in MRI. In: *IPMI 2007*, pp. 713–724 (2007)
7. Fisher, R.A.: The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**(2), 179–188 (1936)
8. McLachlan, G.J.: *Discriminant Analysis and Statistical Pattern Recognition*. Wiley Series in Probability and Mathematical Statistics (1992)

9. Burns, J., et al.: Automated detection of sclerotic metastases in the thoracolumbar spine at CT. *J. Radiol.* **268**(1), 69–78 (2013)
10. Huang, S.F., Chian, K.H.: Automatic detection of bone metastasis in vertebrae by using CT images. In: Proceedings of the World Congress on Engineering, July 2012
11. Yao, J., O'Connor, S.D., Summers, R.M.: Computer aided detection of lytic bone metastases in the spine using routine CT images. In: ISBI, pp. 512–515 (2007)
12. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436 (2015)
13. Roth, H., et al.: Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, vol. 20, pp. 3–12. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-14148-0_1
14. Haider, M.A., et al.: CT texture analysis: a potential tool for prediction of survival in patients with metastatic clear cell carcinoma treated with sunitinib. *Cancer Imaging* **17**(1), 4 (2017)
15. Rao, S.X., et al.: CT texture analysis in colorectal liver metastases: a better way than size and volume measurements to assess response to chemotherapy? *United Eur. Gastroenterol. J.* **4**(2), 257–263 (2016)
16. Vallieres, M., et al.: A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys. Med. Biol.* **60**(14), 5471 (2015)
17. Holli, K., et al.: Texture analysis of MR images of patients with mild traumatic brain injury. *BMC Med. Imaging* **10**(1), 8 (2010)
18. Holli, K., et al.: Characterization of breast cancer types by texture analysis of magnetic resonance images. *Acad. Radiol.* **17**(2), 135–141 (2010)
19. Vinholes, J., Coleman, R., Eastell, R.: Effects of bone metastases on bone metabolism: implications for diagnosis, imaging and assessment of response to cancer treatment. *Cancer Treat. Rev.* **22**(4), 289–331 (1996)
20. Grankvist, J., et al.: MRI and PET/CT of patients with bone metastases from breast carcinoma. *Eur. J. Radiol.* **81**(1), e13–e18 (2012)
21. Bardera, A., Feixas, M., Boada, I.: Normalized similarity measures for medical image registration. In: *Medical Imaging 2004*, pp. 108–118. International Society for Optics and Photonics (2004)
22. Wang, J., et al.: Multi-atlas segmentation of subcortical brain structures via the autoseg software pipeline. *Front. Neuroinformatics* **8**, 7 (2014)
23. Johnson, H., Harris, G., Williams, K., et al.: BRAINSFit: mutual information rigid registrations of whole-brain 3D images. *Insight J.* **57**(1) (2007). <http://hdl.handle.net/1926/1291>
24. Alobaidli, S., et al.: The role of texture analysis in imaging as an outcome predictor and potential tool in radiotherapy treatment planning. *Br. J. Radiol.* **87**(1042), 20140369 (2014)
25. Haidekker, M.: *Advanced Biomedical Image Analysis*. Wiley, Hoboken (2011)
26. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **3**, 610–621 (1973)
27. El Adoui, M., Drisis, S., Larhmam, M.A., Lemort, M., Benjelloun, M.: Breast cancer heterogeneity analysis as index of response to treatment using MRI images: a review. *Imaging Med.* **9**(4), 109–119 (2017)
28. Ojala, T., et al.: A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* **29**(1), 51–59 (1996)

29. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.* **19**(6), 1635–1650 (2010)
30. Liu, D., Wang, S., Huang, D., Deng, G., Zeng, F., Chen, H.: Medical image classification using spatial adjacent histogram based on adaptive local binary patterns. *Comput. Biol. Med.* **72**, 185–200 (2016)
31. Hu, L.S., et al.: Multi-parametric MRI and texture analysis to visualize spatial histologic heterogeneity and tumor extent in glioblastoma. *PLoS ONE* **10**(11), e0141506 (2015)
32. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
33. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* **20**(3), 273–297 (1995)
34. Gressin, A.: Mise à jour d’une base de données d’occupation du sol à grande échelle en milieux naturels à partir d’une image satellite THR. Ph.D. thesis, Université Paris-Descartes (2014)
35. Fisher, R.: The use of multiple measurements in taxonomic problems. *Ann. Hum. Genet.* **7**(2), 179–188 (1936)
36. Cover, T., Hart, P.: Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13**(1), 21–27 (1967)
37. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
38. Hsu, C.W., Chang, C.C., Lin, C.J., et al.: *A Practical Guide to Support Vector Classification* (2003)
39. Raschka, S.: *Python Machine Learning*. Packt Publishing, Birmingham (2015)
40. Varma, S., Simon, R.: Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* **7**(1), 1 (2006)
41. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *14th International Joint Conference on AI* (1995)
42. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46**, 389–422 (2002)
43. Larhman, M.A., Benjelloun, M., Mahmoudi, S.: Vertebra identification using template matching model and k-means clustering. *Int. J. Comput. Assist. Radiol. Surg.* **9**(2), 177–187 (2014)



Parametric Variations of Anisotropic Diffusion and Gaussian High-Pass Filter for NIR Image Preprocessing in Vein Identification

Ayca Kirimtat and Ondrej Krejcar^(✉) 

Faculty of Informatics and Management, Center for Basic and Applied Research,
University of Hradec Kralove, Rokitanskeho 62,
500 03 Hradec Kralove, Czech Republic
a.kirimtat@gmail.com, ondrej@krejcar.org

Abstract. Near infrared (NIR) imaging is one of the promising methods for identification of superficial veins and widely researched and used in clinical medicine and biomedical studies. However, just like imaging in visible spectrum, NIR imaging is not adequate for exact recognition of the vein system as it is, therefore nearly every research starts with preprocessing to prepare the images for identification. Two major filtering methods are anisotropic diffusion and Gaussian high-pass filter which both consist of mandatory parametric adjustments for better visualization of the images and for revealing the vein system. Therefore in this paper we deal with parametric variations of these two methods on a NIR image to give ideas for choosing proper preprocessing techniques and parameters, excluding edge detection and vein detection methodologies.

Keywords: Near infrared · Vein recognition · Preprocessing

1 Introduction

Electromagnetic spectrum, which ranges in between 380 and 700 nm in wavelength, is recognized with human eye and called visible spectrum. After the visible red wavelength, the NIR (near-infrared) spectrum, which ranges from 700 to 1100 nm in wavelength, is located. However, since two of them are located next to each other, there is no correlation between visible and near-infrared spectrum. In addition, due to its hiddenness from a human eye, near-infrared radiation is only captured by special digital cameras. Application areas of near-infrared imaging consist of biometrics, face recognition, biomedicine etc. [1]. The first biomedical application of NIR light, belonged to the work of Cope and Delphy [2], was made to measure blood oxygenation in the brain of babies who were born earlier than the expected time. NIR imaging fundamentally is the non-harmful way for detecting human body [3].

In the spectral band between 300 to 1000 nm in wavelength, NIR illuminator can recognize veins due to the low absorption of skin and subcutaneous tissue. Thus, during NIR lighting, the dark side is mostly blood and the lighter sides are generally fat and skin which provides us to easily recognize veins in hands [4].

Based on the study of Bouzida et al. [4], NIR lighting was used to extract vein pattern to show the high potentiality of the presented method. For the detection, a small portable and monochrome camera with the resolution of 1280×1024 pixel was employed to obtain the vein network on the dorsal hand. Non-invasive NIR imaging provided a better representation of the vein structure in the internal layer of the skin.

Wang et al. [5] investigated two different non-invasive imaging technique explicitly far-infrared imaging and near-infrared imaging. Experiments in various parts of hand were carried out to compare the performances of two imaging methods in thermography. Based on the results of the experiments, near-infrared imaging was more successful than far-infrared imaging for vein pattern biometrics.

In the study of Tien et al. [6], different parts of a hand which are fingers, wrist, palm and arm were detected with NIR imaging using transmission mode, reflection mode and combination of these two modes. The authors used wavelength between 750 and 940 nm to identify veins in higher contrast and they used crossed polarizers and neutral density filters to avoid glare on the detection surface. As a result of their work, both advantages and disadvantages of the presented method were discussed after evaluation and analysis.

Another study by Paquit et al. [7], was conducted for the localization of arm surface veins through near-infrared imaging and structured light ranging technique. The proposed method showed effectiveness in terms of high visibility of surface veins. In the localization phase of this medical procedure, 2D near-infrared images were obtained using standard image processing methods. As a result, authors created 3D topographic map of the arm surface to represent the veins in a better way.

In the article of Alpar and Krejcar [8], anisotropic diffusion and maximum curvature methods were used to extract and estimate hand veins using recorded thermal images. Based on the results, it is highly possible to see temperature differences and hand vein structure through dorsal hand thermograms. In another article of Alpar and Krejcar [9], a novel quantization and indexing algorithm in hand thermography were proposed. They firstly used pseudocolor map and then created grayscale and 4-level quantization after applying red-channel conversion. Based on the histograms and image equalization, the results were quite convincing in terms of contrast differences and hand vein extraction.

As seen from the previous studies, near-infrared imaging is a fast and useful method for the identification of veins in human body. Due to its spectral band, near-infrared lighting techniques are frequently used by medical researchers to extract vein structure in biomedical applications. For the reason that NIR imaging is called a non-invasive method during the clinical phases, this speciality should be used with other filtering methods for conducting more quantitative studies in biomedicine. Therefore, in this article, we mainly focus on two different filtering technique namely Anisotropic Diffusion and Gaussian High Pass Filter to present better vein system visualization for a forearm. In the continuation of this article, Sect. 2 simply explains the Anisotropic Diffusion while Sect. 3 is describing the Gaussian High Pass Filter. Our parametric experiments with these two filtering techniques are carried out for a forearm in Sect. 4. Finally, conclusion remarks and discussion are presented in Sect. 5.

2 Anisotropic Diffusion

Given a grayscale NIR image, any pixel $p_{x,y}$ on the image $R_{x,y}$ could be represented as:

$$p_{x,y} \in R_{x,y}(i = [1 : w], j = [1 : h]) \quad (1)$$

where w is the width, h is the height of the image and x, y are the coordinates of the pixel. If we apply anisotropic diffusion to highlight the vein system on the image, we firstly take partial derivatives by:

$$\frac{\partial R(\bar{x}, \Delta t)}{\partial t} = \nabla(c(\bar{x}, \Delta t) \nabla R(\bar{x}, t)) \quad (2)$$

where, $c(\bar{x}, t)$ is the diffusion function, \bar{x} is the spatial dimensions and Δt the time. In each direction, diffusion is applied N times from the starting point (x, y)

$$R_{n+1}(x, y) = R_n + \frac{c_{d,\Delta t}(x, y) \nabla R_{d,t}(x, y)}{4} \quad (3)$$

where $n = 0, 1, \dots, N$. At this point, there is two option for calculating the diffusion equation $c_{d,\Delta t}(x, y)$ which will be stated as method 1: $m = 1$

$$c_{d,t}(x, y) = e^{\left(\frac{-\nabla(R_{d,\Delta t})}{k}\right)^2} \quad (4)$$

and method 2: $m = 2$

$$c_{d,t}(x, y) = \frac{1}{1 + \left(\frac{-\nabla(R_{d,\Delta t})}{k}\right)^2} \quad (5)$$

where

$$A_{x,y} = R_{d,\Delta t} = R_{\Delta t}(x, y + 1) - R_{\Delta t}(x, y) \quad (6)$$

and k is the gradient threshold. As seen in the equations presented above, there basically are four parameters: N as number of iterations, Δt as integration constant, k as gradient threshold and method m .

3 Gaussian High-Pass Filter

Gaussian high pass filter is rather easy to comprehend and low in detail and parameters; while it is one of the well-known methods in preprocessing of NIR images. It is one of the examples that apply a distribution to the images depending on the distance of the values of the pixels and designated variance. Although the primary utilization area is blurring the images, depending on the selected parameters, it is possible to enhance the

vein system in NIR images. Given the image $R_{x,y}$ as a w-by-h matrix, it is firstly necessary to transform the image by fast Fourier transformation and rearranging the transformation by shifting the zero-frequency component to the center of the arrays.

$$Y_{x+1,y+1} = \sum_{j=0}^{w-1} \sum_{k=0}^{h-1} \omega_w^{jp} \omega_h^{kq} R_{x+1,y+1} \quad (7)$$

where ω is the complex roots of unity represented by $\omega_w = e^{-2\pi i/w}$ and $\omega_h = e^{-2\pi i/h}$. After rearranging the matrix by shifting zero frequency region into the middle, the Gaussian high-pass filter is applied by:

$$G(x, y) = a \left(1 - e^{-d^2(x,y)/2d_0^2} \right) - b \quad (8)$$

where

$$d(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (9)$$

where x_0 and y_0 represent the middle of the image. Afterwards inverse zero frequency shifting and inverse Fourier transformations are successively applied. In this kind of filtering there are three major parameters that can be altered upon requirements: a , b and d_0 . In the following section, each method is separately examined with a NIR image consisting of a forearm.

4 Single Parametric Experiments

On the NIR image of a forearm, we applied both methods with the NIR image itself $R_{x,y}$ and complementary image $R_{x,y}^{-1}$ which could be achieved by $R_{x,y}^{-1} = 255 - R_{x,y}$ for the images in uint8 format which is 8-bit integer or $R_{x,y}^{-1} = 1 - R_{x,y}$ for double precision images, which are all matrix operations. The images we're dealing with are presented in Fig. 1 below.

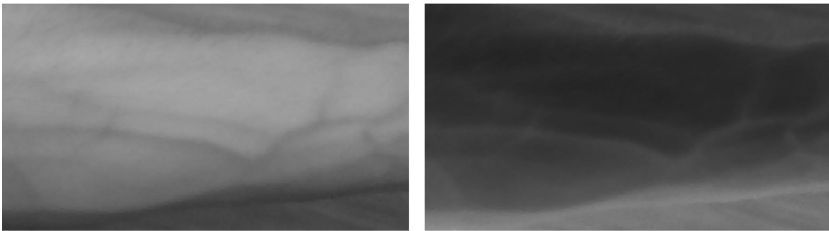


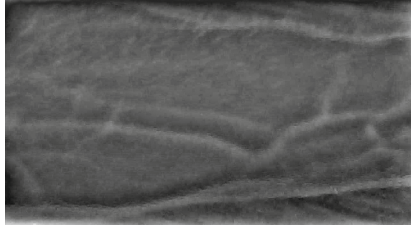
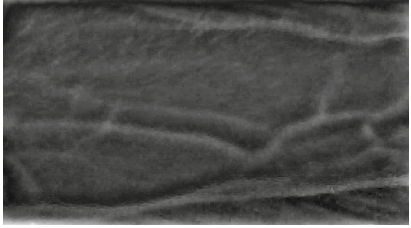
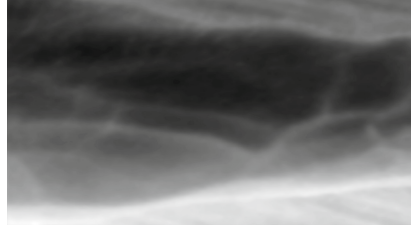

Fig. 1. Original image $R_{x,y}$ on the left and complementary image $R_{x,y}^{-1}$ on the right

Since our original image is in uint8 format, we found the complementary images by subtracting each pixel on the image matrix from 255. As the main difference between these images is the color of the superficial veins: in NIR image the veins system is darker than the forearm; while in complementary image is lighter.

4.1 Gaussian High-Pass Filtering

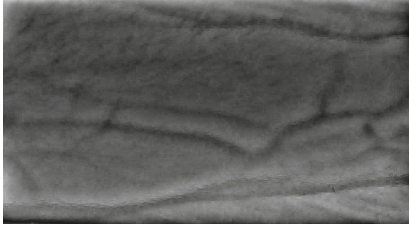


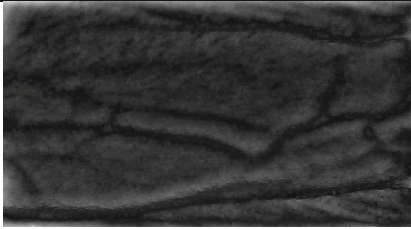
Firstly, we applied Gaussian high pass filter on $R_{x,y}$ by changing the parameters a , b and d_0 . There will be 4 applications presented in Table 1 for this section by altering these parameters. As seen in Table 1, the case with $d_0 = 5$, $a = 200$, $b = 20$ parameters provides better recognition of the veins in contrast with the forearm color. However, in the case with $d_0 = 20$, $a = 150$, $b = 5$ parameters, the veins are unrecognizable as they seem like lost in the dark background color.

Table 1. Gaussian high-pass application on $R_{x,y}$

$d_0=5, a=80, b=10$	$d_0=5, a=200, b=20$
	
$d_0=30, a=20, b=20$	$d_0=20, a=150, b=5$
	

The same procedure is reiterated on the complementary image presented in Table 2 below. The parameters may be different than the first four experiments since we now try to lighten the darker background through applying the filter to the complementary image. The case with the parameters of $d_0 = 5$, $a = 60$, $b = 5$ showed better result than other three experiments since the veins are more detectable in the forearm color. However, the veins seem much thicker than other images. The veins in the case with the parameters of $d_0 = 10$, $a = 80$, $b = 10$ seem very weak to detect easily when we compare it with other experiments.

Table 2. Gaussian high-pass application on $R_{x,y}^{-1}$

$d_0=5, a=50, b=10$	$d_0=10, a=100, b=5$
	
$d_0=10, a=80, b=10$	$d_0=5, a=60, b=5$
	

As seen in Tables 1 and 2, it is possible to state that complementary images give better results in Gaussian high pass filter, mostly in $d_0 = 5, a = 60, b = 5$ case. Because, the contrast difference between the veins and the forearm is more detectable than other three experiments even though the veins are thicker.

4.2 Anisotropic Diffusion Filtering

Secondly, we applied Anisotropic Diffusion Filtering again on both original NIR image and the complementary image separately. Four different parameters explicitly $N, \Delta t, k, m$ belong to the presented method are changing to present various visualization for the veins in the same forearm.

In Table 3 below, the different parameters of Anisotropic Diffusion Filtering applied to the original image are shown with the images. The veins in the case with the parameters of $N = 50, \Delta t = 0.1, k = 30, m = 2$ are so weak in contrast with the lighter forearm color which means it is difficult for us to recognize the veins when we compare it with other three experiments. On the other hand, the contrast difference between the veins and the forearm in the case with parameters of $N = 20, \Delta t = 0.025, k = 20, m = 1$ is quite better among other experiments.

According to Table 4 below, the different parameters of Anisotropic Diffusion Filtering were used on the complementary image to again present various images. The veins in the case with the parameters of $N = 80, \Delta t = 0.02, k = 50, m = 1$ are so weak to detect also due to the light gray forearm color. However, the veins in the case with the parameters of $N = 50, \Delta t = 0.02, k = 25, m = 1$ is best recognizable among other

images since the forearm color is darker than other three images. As a result, Anisotropic Diffusion Filtering gives better results on the complementary image rather than original NIR image in terms of color contrasts.

Table 3. Anisotropic diffusion application on $R_{x,y}$

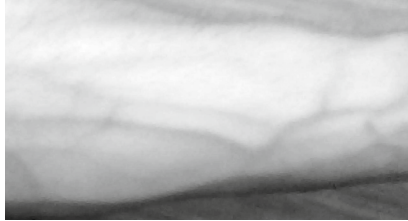







N=20, $\Delta t = 0.025$, $k=20$, $m=1$	N=50, $\Delta t = 0.1$, $k=30$, $m=2$
	
N=30, $\Delta t = 0.2$, $k=10$, $m=2$	N=80, $\Delta t = 0.005$, $k=25$, $m=1$
	

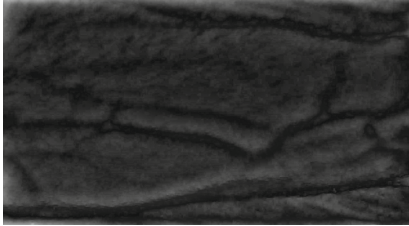
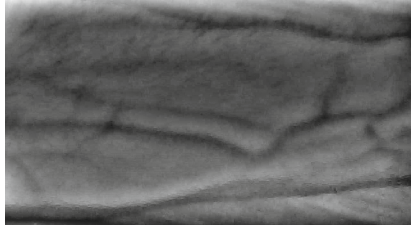
Table 4. Anisotropic diffusion application on $R_{x,y}^{-1}$

N=80, $\Delta t = 0.02$, $k=50$, $m=1$	N=50, $\Delta t = 0.02$, $k=25$, $m=1$
	
N=40, $\Delta t = 0.001$, $k=10$, $m=2$	N=100, $\Delta t = 0.01$, $k=100$, $m=2$
	

4.3 Combined Filtering

Finally we tried a combined filtering starting with anisotropic diffusion and subsequently Gaussian high-pass filter on the complementary image $R_{x,y}^{-1}$. Table 5 contains two different images which are performed with two filtering methods and various parameters respectively.

Table 5. Combined filtering on $R_{x,y}^{-1}$

N=20, $\Delta t = 0.05$, $k=20$, $m=2$	N=10, $\Delta t = 0.05$, $k=25$, $m=1$
$d_0=5$, $a=60$, $b=5$	$d_0=3$, $a=75$, $b=20$
	

As seen in the combined filtering, the case with the parameters $N = 20$, $\Delta t = 0.05$, $k = 20$, $m = 2$ and $d_0 = 5$, $a = 60$, $b = 5$ presented better results than the other experiment, since we can easily recognize the veins in contrast with the forearm color. Although the other case on the right can be better than other experiments in the previous sections, however, it is not easy do detect the veins, because of the lighter forearm color and unrecognizable and blurred vein contours.

5 Conclusion and Discussion

Near-infrared imaging is one of the strongest preprocessing method for revealing the vein structure of any part of the human body. In addition, it is a well-known technique in biomedical applications since it does not give any radiation to human body. Under the skin layer, near-infrared lights can easily penetrate into inner layers such as veins in the forearm or hands. However, after recording the vein system of the human body with special near-infrared cameras, filtering and image processing should be implemented to better recognize the contrast between the background and the vein structure. In existing literature, there are many filtering methods used in the biomedical applications for approximately 30 years.

In this article, we apply two major filtering techniques namely Anisotropic Diffusion Filtering and Gaussian High-Pass Filtering to both original NIR image and complementary image separately. When we change each parameter for both filtering method, the results give a different image each time. The best results are obtained using combination of these two methods to the complementary image since we can easily recognize the darker vein structure of the gray forearm color.

As seen from the image results in the experiments, the worst cases are seen in the application of Anisotropic Diffusion Filtering on the original image. Because, the vein system is not visible enough in comparison with the lighter forearm color. Besides, the case with the parameters of $d_0 = 20$, $a = 150$, $b = 5$ results in unrecognizable vein structure on the darker forearm color. On the other hand, the best result belongs to the combined filtering on the complementary image with the parameters of $N = 20$, $\Delta t = 0.05$, $k = 20$, $m = 2$ and $d_0 = 5$, $a = 60$, $b = 5$ since the vein system is dark and thicker in contrast with the forearm color.

Since the NIR imaging is frequently used method by clinical researchers, it should be enhanced with such filtering methods during the clinical phases for quantitative detection of the medical problems, vein structure and blood flow in biomedicine. Thus, we presented promising results in terms of vein structure extraction using major filtering methods. Moreover, we showed the high potentiality of NIR imaging as a preprocessing method before better visualization of the vein structure. More future studies should be conducted to use the valuable utility of NIR imaging in the biomedical applications around the world. Besides, both quantitative and qualitative methods should be performed together especially in the biomedicine.


Acknowledgement. The work and the contribution were supported by the SPEV project “Smart Solutions in Ubiquitous Computing Environments”, University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic (under ID: UHK-FIM-SPEV-2018).

References

1. Fredembach, C.: Near-infrared imaging. <http://ivrl.epfl.ch/research/infrared/imaging>. Accessed 14 Jan 2018
2. Cope, M., Delpy, D.T.: System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Med. Biol. Eng. Comput.* **26**(3), 289–294 (1988)
3. Xueyan, L., Shuxu, G.: The fourth biometric - vein recognition. In: Yin, P.-Y. (ed.) *Pattern Recognition Techniques, Technology and Applications* (2008)
4. Bouzida, N., Bendada, A.H., Maldague, X.P.: Near-infrared image formation and processing for the extraction of hand veins. *J. Mod. Opt.* **57**(18), 1731–1737 (2010)
5. Wang, L., Leedham, G., Cho, S.Y.: Infrared imaging of hand vein patterns for biometric purposes. *IET Comput. Vis.* **1**(3), 113–122 (2007)
6. Van Tien, T., Mien, P.T.H., Dung, P.T., Linh, H.Q.: Using near-infrared technique for vein imaging. In: Toi, V.V., Lien Phuong, T.H. (eds.) *5th International Conference on Biomedical Engineering in Vietnam*. IP, vol. 46, pp. 190–193. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-11776-8_46
7. Paquit, V., Price, J.R., Seulin, R., Mériaudeau, F., Farahi, R.H., Tobin, K.W., Ferrell, T.L.: Near-infrared imaging and structured light ranging for automatic catheter insertion. In: *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display* (2006)
8. Alpar, O., Krejcar, O.: Superficial dorsal hand vein estimation. In: Rojas, I., Ortuño, F. (eds.) *IWBBIO 2017. LNCS*, vol. 10208, pp. 408–418. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56148-6_36
9. Alpar, O., Krejcar, O.: Quantization and equalization of pseudocolor images in hand thermography. In: Rojas, I., Ortuño, F. (eds.) *IWBBIO 2017. LNCS*, vol. 10208, pp. 397–407. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56148-6_35



FLIR vs SEEK in Biomedical Applications of Infrared Thermography

Ayca Kirimtat and Ondrej Krejcar^(✉) 

Faculty of Informatics and Management, Center for Basic and Applied Research,
University of Hradec Kralove, Rokitanskeho 62,
500 03 Hradec Kralove, Czech Republic
a.kirimtat@gmail.com, ondrej.krejcar@uhk.cz

Abstract. This article aims to compare the results of two different infrared cameras through revealing the irregular temperature distribution on an injured toe. Since the inhomogeneous body temperature is the key indicator of severe injuries, wounds, and illnesses, infrared thermography is the strongest method among other conventional methods to map the skin temperature variations. The current utility of infrared cameras in biomedical applications is also presented to comprehend the relationship between the nature of thermal radiation and human body temperature. In the article, the presented biomedical applications include skin cancer screening, wound detection in a diabetic foot, muscle activation assessment during an exercise, or thermal mapping of healthy human bodies. Along with the developments in infrared thermography, this article focuses on analyzing temperature distribution on the injured toe of a subject with two different smartphone-based infrared camera models namely FLIR One and SEEK Compact Pro. In addition, the results obtained from the presented infrared camera models are compared regarding to the thermal images of the injured toe.

Keywords: Infrared thermography · Biomedicine · Thermal camera
FLIR · SEEK

1 Introduction

Along with the technological developments, standard cameras, that capture only visible light in RGB or grayscale, become inadequate to recognize both human and objects in totally dark environments. Over time, passive sensors integrated to infrared cameras emerged to recognize the objects between the 3–14 μm infrared spectrums. Therefore, in many applications such as biomedicine, transportation, fire safety and buildings, infrared cameras are preferred mostly, since they do not generally depend on additional heat sources for making thermal detection [1]. In addition, both data collection and processing is highly achievable with infrared cameras since they create real time and high resolution thermal images. Thus, thermography experiments should be carried out with infrared cameras rather than standard thermometers in biomedical applications where temperature changes uncontrollably [2].

Basically, infrared cameras measure the emitted radiation, a function of surface temperature, from an object surface and turn it to thermal images. Thus, the camera can

simply calculate and reveal inhomogeneous temperature differences [3]. The infrared cameras in the current commercial market are the third generation thermal cameras which have FPA (focal plane array) detectors. FPA detectors mainly consist of photon detectors (cooled) and thermal detectors (uncooled) [4]. Cooled detectors are generally used for scientific or military purposes, while uncooled detectors are used for transportation, fire safety, building or biomedical applications. FLIR and SEEK infrared camera models, which are used for the purpose of biomedical detection, mostly work with thermal detectors. According to the literature survey, FLIR infrared camera models are preferred more than smartphone-based SEEK camera models in biomedicine, since it also works with photon detectors together with thermal detectors.

FLIR infrared cameras are generally standalone camera models which offer accurate, quantitative and in-depth measurements. FLIR cameras reveal temperature variations even under $0.02\text{ }^{\circ}\text{C}$ (20 mK) which means the deepest body or skin detection is highly possible in biomedical applications. Through the high potentiality of FLIR cameras in biomedicine, medical researchers can reach more informative results [5]. On the other hand, even though SEEK thermal cameras work with thermal detectors which are appropriate to record thermal images in biomedicine, they are mostly used for building diagnostics, outdoors, firefighting and commercial trades. Because, they are mostly designed for finding elementary heat sources in smartphone-based applications which can turn a smartphone into a thermal camera. SEEK thermal cameras are also preferred since they are generally more economical than other thermal camera models [6]. Table 1 summarizes the overall differences and similarities of two infrared camera models.

Table 1. FLIR vs SEEK model infrared cameras.

	FLIR infrared camera	SEEK infrared camera
Application area	Buildings, biomedicine, industry, lab tests	Buildings, outdoors, commercial, firefighting
Detector types	Thermal and photon detectors	Thermal detector
Generation	3 rd generation	3 rd generation
Infrared spectrums	3–14 μm	3–14 μm
Portability	Standalone and smartphone-based compact	Smartphone-based compact
Measurement type	Quantitative and qualitative	Qualitative

2 FLIR and SEEK Infrared Cameras in Biomedicine

IR thermal detection for medical purpose is both passive and non-invasive method for revealing temperature differences in body, skin and organs. Rather than body structure, IR thermography is mostly used for the body shell inspections in biomedicine. Once the thermal detection is made with high resolution by the support of image processing with computer software, reliable results with huge data acquisition is quite inevitable. Another advantage of IR thermography is being lack of radiation which is a valuable factor in biomedical applications [7]. Based on the medical literature, several

applications with FLIR and SEEK infrared cameras can be found to conduct thermographic surveys in biomedicine.

Contreras et al. [8] used an infrared device namely FLIR E60 IR camera to obtain thermal radiation of the feet of the participants. All participants were asked to remove everything from their feet in a conditioned room. In order to avoid from the radiation of the body, the IR camera was positioned 1 m away from the feet. According to the results of the experiments, presented methodology was found both useful and accurate for making such an assessment in the diabetic feet. In addition, the method provided valuable information, since it is quantitative and comprehensive.

Another study by Quesada et al. [9] was conducted to analyze the relationship between participants' neuro-muscular activation and skin temperature during a cycling test. Thermal images were recorded both in the end of the test and 10 min after the test using FLIR E60 IR camera in a controlled environment. For the thermographic survey, 10 sportive people voluntarily participated in the laboratory tests. They were asked not to drink alcohol, smoke and other stimulants such as coffee 12 h before the tests. Results of tests showed that studying all differences between trained and untrained participants is still lack in the literature. However, the participants, who showed higher overall activation and lower vastus lateralis activation, adaptively response better than other participants and they also had limitation in the increase of their skin temperature during the tests.

In the study of Fournet et al. [10] body-mapping approach was used for 18–25 aged 9 males and 9 females to investigate regional body temperature. Their whole-body temperature was recorded using an infrared camera namely Thermo-cam B2, FLIR Systems. The presented approach mainly aimed to find spatial thermoregulatory and perceptual responses of each body. The experiments started with usual outdoor running activities and continued with a rest in a conditioned room as four stages. Regarding to the running results, female bodies showed lower skin temperatures than male bodies even they both exhibited similar overall thermal sensations.

Chudecka and Lubkowska [11] used a thermal imaging camera (ThermaCAM SC500 FLIR System), which is used for a valuable diagnosis for physiotherapy, medicine and sport disciplines, to reveal thermal maps of young women and men. Temperature distribution on their bodies was also examined based on participants' ages, body surface areas and fat distribution. All of the sportive participants were healthy with no injuries on their bodies. The research demonstrated that there is a quite difference in the body temperatures of men and women.

According to Bernard et al. [12], using IR imaging in biomedicine and scientific studies is unavoidable, since it is a unique tool for the detection of object temperature. Thus, a thermographic inspection was carried out with FLIR B200 thermal camera (equipped with FPA microbolometer thermal sensor) to investigate whether physical context of human body affect the body temperature or not. The authors checked the skin surface temperature with the infrared camera after applying ultrasound gel, water and hydration cream on the patients' hands. As a result, they found IR cameras quite useful for safe and contactless measurements in biomedical applications by also taking into consideration the effect of emissivity values.

Alpar and Krejcar [13, 14] presented two different studies in IWBBIO 2017 conference about hand thermography. Smartphone-based FLIR One was used for thermal

imaging in both studies. Based on the studies, hand thermograms are remarkably helpful and informative to recognize hand veins by creating temperature differences. In addition, the results obtained from the studies were promising enough to create contrast differences.

On the other hand, existing literature has lack of SEEK infrared camera usage in biomedical applications since it is more suitable for being used in other areas such as firefighting, animals or buildings. Even it mostly works with smartphone-based applications, Peleki and da Silva [15] tried portable SEEK Thermal Compact XR for the immediate detection of Acute Limb Ischaemia. The authors discussed its applicability on the assessment of tissue perfusion when a 78-year-old woman was in the bed during clinical evaluation of her illness. As a conclusion of their research, the presented device, which is compatible with many smartphone models, provided baseline information before main evaluations and controls. Future studies should be carried out on the applicability of the device for quantitative assessments in biomedicine.

In this article, we aim to compare the resulting thermograms of injured toe of a subject using two different smartphone-based thermal cameras namely FLIR One and SEEK Compact PRO. Through recording images with two infrared cameras, various formats are applied to the images to obtain visible results in the thermograms. In the continuation of this research, Sect. 3 defines the similarities and differences of the presented infrared cameras. The experiments conducted with two infrared cameras are included in detail in Sect. 4. The discussion part and the conclusion remarks are given in Sect. 5.

3 Main Differences of FLIR One and SEEK Compact PRO

In our biomedical research, we mainly use FLIR One and SEEK Compact PRO thermal cameras as shown in Fig. 1 and they are manufactured for Android devices with their designated applications. Table 2 simply summarizes the main differences of the presented thermal cameras. The first difference is charging necessity of FLIR Camera since it has an integrated battery that needs to be recharged very frequently; while SEEK camera takes the energy from the android device which is definitely more user friendly.



Fig. 1. The cameras we benchmarked in this paper: SEEK Compact PRO and FLIR One

Table 2. FLIR One vs SEEK Compact Pro.

	FLIR One	SEEK Compact PRO
Operated device	Android	Android
Detector types	Thermal and photon detectors	Thermal detector
Generation	3 rd generation	3 rd generation
Infrared spectrums	3–14 μm	3–14 μm
Portability	Smartphone-based compact	Smartphone-based compact
Image extraction format	Own format in a database	JPEG
Image resolution	343 \times 458 portrait	1280 \times 720 landscape
Battery usage	Own battery	Battery of the android device

Second difference is the wide but scrambled spectrum of FLIR camera that is said to have thermal and photon detectors together which is totally correct. If there is a substantial illumination, the images consist of two edges which are noisier than SEEK, yet the image quality is significantly higher. Therefore, the experiments with FLIR should be conducted where there is no direct light on the object or no dominant light condition.

The other differences of usability and technical specifications could be summarized as: FLIR has a converter to change the angle of the camera which could be extremely important for some cases. SEEK's image extraction procedure is very easier than FLIR since the images are in JPEG format stored in a separate folder; while FLIR saves the images in its own format in a database. Although it seems rather complicated, it gives users the opportunity of altering the pseudocolor of a thermogram after it is captured.

As we only focus on biomedical applications, not industrial, both cameras provide a proper information of the temperature when a point of the screen is selected. The basic image format of SEEK is landscape in 1280 \times 720 resolution; while FLIR extracts portrait images in 343 \times 458. Each camera offers dynamic contrasting, very similar pseudocolor portfolio, and plug-n-play simplicity.

4 Experiments

The experiment is on the injured toe of a subject 35/M when he is on convalescence period. Under the same light and temperature conditions and without footwear, experiments are conducted by SEEK Compact PRO and FLIR One separately and the resulting thermograms are captured with original grayscale and various pseudocolored formats. The difference in image quality is also seen in all figures which means thermograms captured by FLIR One have higher image quality than SEEK Compact PRO even the thermal images are noisy in FLIR One.

4.1 Experiments with SEEK Compact PRO

The first experiment is focusing on the toes of the subject separately with built-in grayscale thermograms. All thermograms belong to SEEK Compact PRO are in

1280 × 720 landscape image resolution as seen in figures below. If we compare one by one, due to dynamic filtering, it is not possible to identify whether there is a medical problem with the subject since right foot seems to be hotter than the left one while the background is in different color. Figure 2 shows the grayscale thermograms captured by SEEK Compact PRO separately for both foot.

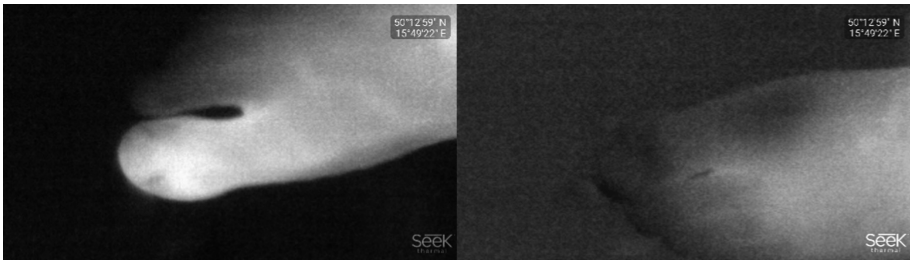


Fig. 2. The grayscale thermograms for each foot captured by SEEK Compact PRO

Therefore, we turned into rainbow pseudocolor to find a way to discriminate the features of contrast, however the hot and cold places of the thermogram is again readjusted by dynamic contrasting. Two feet were again captured separately in rainbow pseudocolor format. As it is seen from Fig. 3, red color represents the hot places in both thermogram. The toe in red color on the right foot refers to injured part.

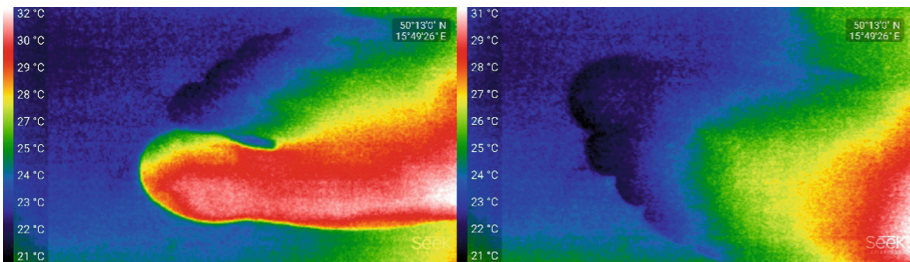


Fig. 3. The rainbow colored for each foot captures by SEEK Compact PRO (Color figure online)

For understanding the performance of contrast formation in pseudocolored thermograms, both feet are captured together with various formats. Therefore, the difference between the temperatures is totally visible. Figure 4 shows four different pseudocolored thermograms of SEEK for both foot. In the bottom part of Fig. 4, both format type in the thermograms originally belong to the SEEK camera which means they do not exist in FLIR camera formats.

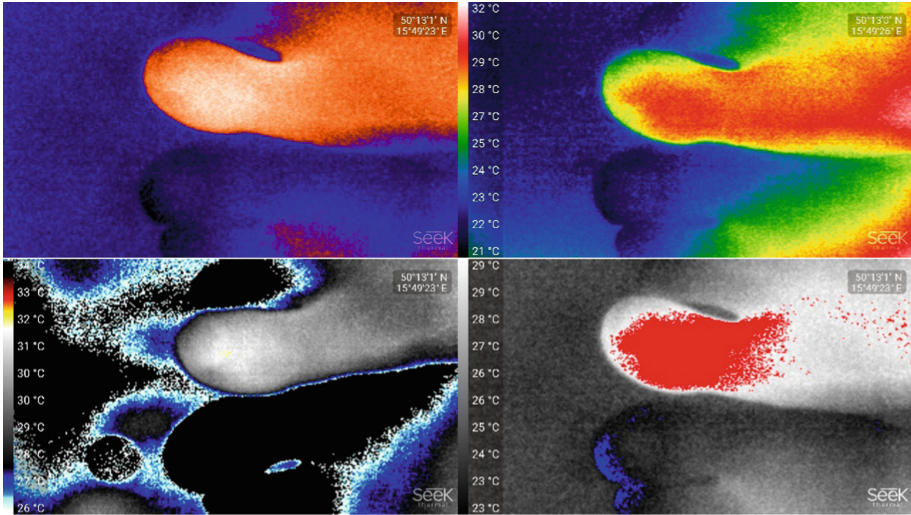


Fig. 4. Various pseudocolored thermograms of SEEK Compact PRO for both foot.

4.2 Experiments with FLIR One

The second experiment was carried out with FLIR One camera under again the same conditions and on the same subject both separately and together. The thermograms captured with FLIR One are in 343×458 portrait and noisy. Figure 5 presents the grayscale thermograms of right and left foot captured separately. One by one comparison is again less satisfied than simultaneous comparison. Grayscale thermograms in Fig. 6 are better recognized in terms of color contrasts which refer to better temperature differences.



Fig. 5. The grayscale thermograms for each foot captured by FLIR One



Fig. 6. The grayscale thermogram for both foot captured by FLIR One

The pseudocolored thermograms of FLIR One in Fig. 7 originally belong to the camera itself since they are intentionally captured to show noticeable temperature. In both thermograms, we can recognize the injured toe in the right foot easily, since it is in different colors in both thermogram. In addition, all toes in the left foot are in blue and gray which means they represent the cold parts while hot parts are in red. As it is

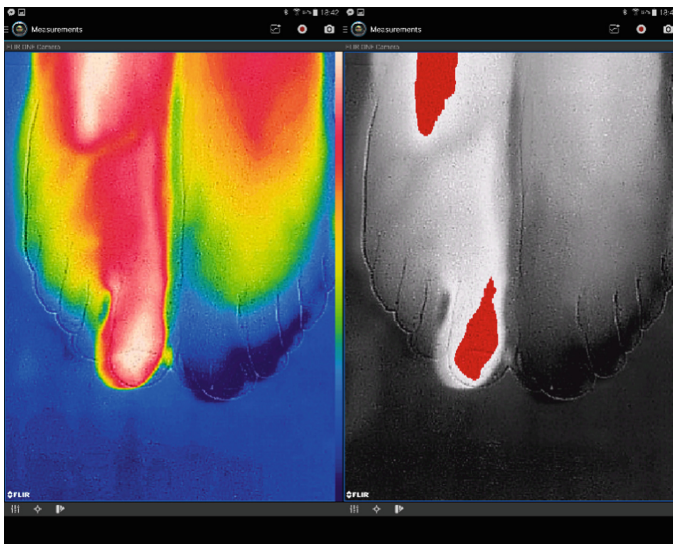


Fig. 7. Various pseudocolored thermograms of FLIR One for both foot.

clearly seen from all the thermograms belong to FLIR One that, contours in the edges of the feet and toes exist and this speciality is available in FLIR One camera rather than SEEK Compact PRO.

5 Conclusion and Discussion

Two different smartphone-based infrared camera models are used to obtain digital thermograms after applying various formats to the thermograms of the injured toe in the right foot by making a comparison with the left toes of the subject. Both SEEK Compact PRO and FLIR One are quite suitable for close-up shots for this kind of biomedical application. However, as it is seen from the figures in Sect. 4, close-up shots with SEEK Compact PRO are considerably better than close-up shots with FLIR One. However, one by one dynamic filtering in grayscale format with SEEK Compact PRO is not pleasing to detect the medical problem on the injured toe. Because, the subject is on convalescence period and contrast difference between background and the subject is not satisfactory enough. Even the comparison is again made one by one, hot and cold regions are easily perceived with dynamic contrasting after applying rainbow pseudocolor format. To create more detectable contrast difference, two feet are captured together in various pseudocolored thermograms of SEEK Compact PRO.

On the other hand, due to high image quality and visibility of the subject contours in the FLIR One camera, the contrast difference between the subject and the background is definitely better. Various pseudocolored thermograms are also obtained with FLIR One to present different format performances. The pseudocolored thermograms of FLIR One originally belong to the camera since it is intentionally captured to show noticeable temperature. The main disadvantage of the thermograms with FLIR One is that the images are noisy.

As a consequence, traditional cameras and thermometers fail to satisfy detection in many areas along with the technological developments that supply infrared cameras to the industry. Several applications of infrared thermography such as transportation, buildings, firefighting and biomedicine etc. can be found in existing literature. However, infrared cameras allow more precise and non-invasive observation of medical problems rather than conventional methods in biomedicine. In addition, during the clinical assessment of the patients, the subject is not exposed to unsafe radiation which is a unique advantage of the technique. Easy and quick evaluation of even simple medical problems can be provided through smartphone-based thermal cameras.

The main contribution of the study is to provide several thermograms of a medical problem belonged to two specific smartphone-based infrared cameras and to show effectiveness of the presented method for quick and comparative diagnosis in biomedicine. Additionally, based on the similar and different characteristics of the presented cameras given in the previous sections, appropriate thermal camera type could be preferred for an ideal and noticeable biomedical detection. Together with the comparative diagnosis, more future experiments should be carried out in the context of quantitative diagnosis of biomedical problems to explore high potentiality of the presented infrared cameras.

Acknowledgement. The work and the contribution were supported by the SPEV project “Smart Solutions in Ubiquitous Computing Environments”, University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic (under ID: UHK-FIM-SPEV-2018).

References

1. Gade, R., Moeslund, T.B.: Thermal cameras and applications: a survey. *Mach. Vis. Appl.* **25** (1), 245–262 (2014)
2. Lahiri, B.B., Bagavathiappan, S., Jayakumar, T., Philip, J.: Medical applications of infrared thermography: a review. *Infrared Phys. Technol.* **55**, 221–235 (2012)
3. Buzug, T.M., Schumann, S., Pfaffman, L., Reinhold, U., Ruhlmann J.: Functional infrared imaging for skin-cancer screening. In: *Proceedings of 28th IEEE EMBS Annual International Conference*, New York City, USA, pp. 2766–2769 (2006)
4. Cardone, D., Merla, A.: New frontiers for applications of thermal infrared imaging devices: computational psychophysiology in the neurosciences. *Sensors* **17**, 1042 (2017)
5. FLIR Research and Science. <http://www.flir.com/science/content/?id=78846>. Accessed 11 Jan 2018
6. SEEK Thermal Homepage. <https://www.thermal.com/compact-series.html>. Accessed 11 Jan 2018
7. Szentkuti, A., Kavanagh, H.S., Grazio, S.: Infrared thermography and image analysis for biomedical use. *Periodic. Biolog.* **113**(4), 385–392 (2011)
8. Contreras, H.D., Barreto, H.P., Magdaleno, J.R., Bernal, J.A.G., Robles, L.A.: A quantitative index for classification of plantar thermal changes in the diabetic foot. *Infrared Phys. Technol.* **81**, 242–249 (2017)
9. Quesada, J.I.P., Carpes, F.P., Bini, R.R., Palmer, R.S., Soriano, P.P., De Anda, R.M.C.O.: Relationship between skin temperature and muscle activation during incremental cycle exercise. *J. Therm. Biol.* **48**, 28–35 (2015)
10. Fournet, D., Ross, L., Voelcker, T., Redortier, B., Havenith, G.: Body mapping of thermoregulatory and perceptual responses of males and females running in the cold. *J. Therm. Biol.* **38**, 339–344 (2013)
11. Chudecka, M., Lubkowska, A.: Thermal maps of young women and men. *Infrared Phys. Technol.* **69**, 81–87 (2015)
12. Bernard, V., Staffa, E., Mornstein, V., Bourek, A.: Infrared camera assessment of skin surface temperature e effect of emissivity. *Physica Med.* **29**, 583–591 (2013)
13. Alpar, O., Krejcar, O.: Superficial dorsal hand vein estimation. In: Rojas, I., Ortuño, F. (eds.) *IWBBIO 2017. LNCS*, vol. 10208, pp. 408–418. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56148-6_36
14. Alpar, O., Krejcar, O.: Quantization and equalization of pseudocolor images in hand thermography. In: Rojas, I., Ortuño, F. (eds.) *IWBBIO 2017. LNCS*, vol. 10208, pp. 397–407. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56148-6_35
15. Peleki, A., da Silva, A.: Novel use of smartphone-based infrared imaging in the detection of acute limb ischaemia. *EJVES Short Rep.* **32**, 1–3 (2016)



Advances in Homotopy Applied to Object Deformation

Jose Alejandro Salazar-Castro^{1,2(✉)}, Ana Cristina Umaquina-Criollo^{3,4},
Lilian Dayana Cruz-Cruz², Luis Omar Alpala-Alpala⁵,
Catalina González-Castaño⁶, Miguel A. Becerra-Botero⁷,
Diego Hernán Peluffo-Ordóñez⁸, and Cesar Germán Castellanos-Domínguez¹

¹ Universidad Nacional de Colombia, sede Manizales, Manizales, Colombia
joasalazarca@unal.edu.co

² Corporación Universitaria Autónoma de Nariño, Pasto, Colombia

³ Universidad Técnica del Norte, Ibarra, Ecuador

⁴ Universidad de Salamanca, Salamanca, Spain

⁵ Universidad Cooperativa de Colombia, Pasto, Colombia

⁶ Universitat Rovira i Virgili, Tarragona, Spain

⁷ Institución Universitaria Pascual Bravo, Medellín, Colombia

⁸ Universidad Yachay Tech, Urcuquí, Ecuador

Abstract. This work explores novel alternatives to conventional linear homotopy to enhance the quality of resulting transitions from object deformation applications. Studied/introduced approaches extend the linear mapping to other representations that provides smooth transitions when deforming objects while homotopy conditions are fulfilled. Such homotopy approaches are based on transcendental functions (TFH) in both simple and parametric versions. As well, we propose a variant of an existing quality indicator based on the ratio between the coefficients curve of resultant homotopy and that of a less-realistic, reference homotopy. Experimental results depict the effect of proposed TFH approaches regarding its usability and benefit for interpolating images formed by homotopic objects with smooth changes.

Keywords: Homotopy · Object deformation
Transcendental functions · Smooth transitions

1 Introduction

The process of mapping, by which a continuous function is deformed onto another, is generally known as homotopy [1], and consists of convergent projection or prolongation of functions in different topological spaces. Two objects can be called homotopic when they are continuously deformed in their respective topological spaces [2]. Typically, homotopy has been used as an extension of traditional methods of disturbance, which allows for robust optimization by reducing the complexity of the global system [3]. Different approaches converge in the

usefulness of homotopy to achieve optimization applicable to any problem, such as: the resolution of time-dependent differential equations [4] and search for subdivisions in objects that locally represent a global system [5], among others. As well, homotopy-based methods have been used for different practical applications related to data visualization, modeling and industrial applications [6–8]. Homotopic mapping in its conventional definition is based on polynomial approaches, which according to their nature do not always adequately and realistically represent the deformation of the object. In this connection, this paper introduces a novel approach here called transcendental-function-based homotopy (TFH), which uses transcendental functions as an alternative to conventional homotopic functions. Thus, more realistic transitions are achieved during the deforming process. In other words, TFH, given its essential characteristics, can naturally generate a smoother or gradual transition when one function is transformed into another. Particularly, for object deformation application wherein visual effect is decisive, the smoothness property provided by TFH is certainly of great benefit. Experiments were conducted using a spherical shell that is deformed into a cylinder. Also, images from the Yale Faces database were used [9], were also used to transcribe a smiling face. In addition, a measure of quality related to the area under the curve is used to quantify the performance of the assessed homotopic approaches. The results clearly illustrate the applicability of the proposed homotopy to deform objects within a smooth and dynamic framework.

The remaining of this paper is organized as follows: Sect. 2 describes conventional homotopy based on a linear approach and its polynomial extension, as well as the use of exponential mapping. Section 3 outlines the proposed far-reaching homotopic approach in which the smoothness effect of mapping is analyzed through parametric exponential and sine/cosine functions. Section 4 is an extension of a quality measure. Section 5 holds the experimental setup. Results and discussion are gathered in Sect. 6. Finally, concluding remarks are drawn in Sect. 7.

2 Conventional Homotopy

Homotopy is associated with the concept of disturbance, by including a parameter to make changes on functions that belong to two different topological spaces [3]. This change is made through the mapping from a continuous function to another one, generating successive transformations in their respective topological spaces while preserving their properties, and varying a single homotopic parameter [1]. In Eq. 1, a mathematical form is presented to explain conventional homotopy. Let f_1 and f_2 be two continuous function, then the homotopy mapping h from a space χ to a space ψ is defined as follows:

$$h : \chi \times [0, 1] \rightarrow \psi, \quad f_1, f_2, \lambda \mapsto h(f_1, f_2 | \lambda), \quad (1)$$

such that the $h(f_1, f_2, 0) = f_1$ and $h(f_1, f_2, 1) = f_2$ conditions are fulfilled, where the homotopic parameter is λ . A linear combination is a feasible way to write a homotopy mapping, so that the coefficients are given in terms of λ ,

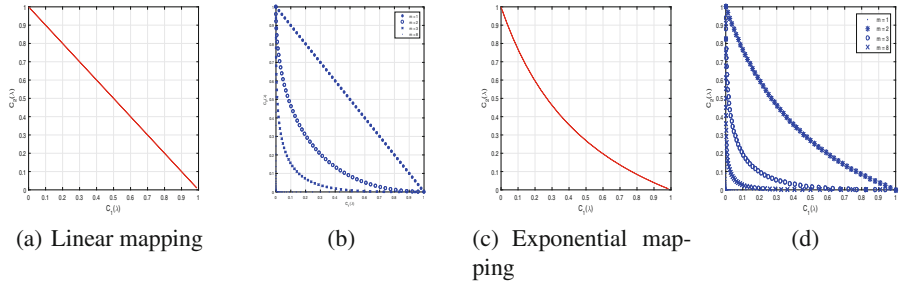


Fig. 1. Coefficients relationship in homotopic mapping and their $C_1(\lambda)$ vs $C_2(\lambda)$ plotting at different levels of m . The linear relationship in Fig. 1(a) with its levels of m in Fig. 1(b), and the exponential relationship in Fig. 1(c) with its level of m in Fig. 1(d)

so $h(f_1, f_2, \lambda) = C_1(\lambda) f_1 + C_2(\lambda) f_2$. If the conditions are fulfilled, the coefficients obtained by mapping h become valid. A general homotopy axiom can be expressed for any homotopy function h and any map $h_o : \chi \rightarrow Z$, so there is a map $\tilde{h} : \chi \times [0, 1] \rightarrow \psi$, $h_o : \lambda \mapsto f_1$, $h_o \in \chi, f_1 \in \psi, f_1 \in \chi$, and $f_2 \in Z$ if there exists a mapping $\pi : Z \rightarrow \psi$ for lifting $h_o = h|_{\chi \times \{0\}}$ (i.e., so that $h_o = \pi \circ h$). This axiom should also be fulfilled.

2.1 Linear Mapping

In terms of the coefficients from $h(f_1, f_2, \lambda)$, the coefficients $C_1^{lin}(\lambda) = \lambda$ and $C_2^{lin}(\lambda) = 1 - \lambda$ can express the simplest homotopy. This approach is called linear homotopy [1]. Figure 1(a) shows the inverse relationship between C_1 and C_2 .

2.2 Exponential Mapping

In this approach, followed from a non-linear transformation, the coefficients are given by $C_1^{exp}(\lambda) = e^\lambda$ and $C_2^{exp}(\lambda) = e^{(1-\lambda)}$. These coefficients do not fulfill directly the homotopy conditions. In [1], a linear transformation is introduced in the form $T(x) = ax + b$, being $a = \frac{1}{e-1}$ and $b = \frac{-1}{e-1}$. Doing so, the aforementioned exponential mapping becomes a valid homotopy approach, while normalized coefficients are $T(e^\lambda) = \frac{1-e^\lambda}{1-e} = C_1^{exp}(\lambda)$ and $T(e^\lambda) = \frac{1-e^{(1-\lambda)}}{1-e} = C_2^{exp}(\lambda)$. Again, the inverse relationship between coefficients is verified as depicted in Fig. 1(c).

2.3 Polinomial Mapping

A polynomial approach introduced in [1] allows to extend the linear mapping to a one-degree-of-freedom version as $C_1^{pol}(\lambda) = \lambda^m$ and $C_2^{pol}(\lambda) = (1 - \lambda)^m$, and a generalized exponential homotopy with coefficients in the form $C_1^{exp}(\lambda) = \left(\frac{e^\lambda}{1-e}\right)^m$ and $C_2^{exp}(\lambda) = \left(\frac{e^{(1-\lambda)}}{1-e}\right)^m$, where $m \in \mathbb{R}^+$. The conditions given in Sect. 2 are fulfilled at any m in both linear and exponential extended mapping.

3 Parametric and Trascendental Functions-Based Mapping

In this section, we introduce novel approaches able to improve the deformation performing a “soft” mapping and reaching more realistic deformations. Figure 2 depicts that coefficients are ruled by an inverse relationship fulfilling the conditions given in Sect. 2. All of the proposed mapping approaches allow for a smooth transition behavior.

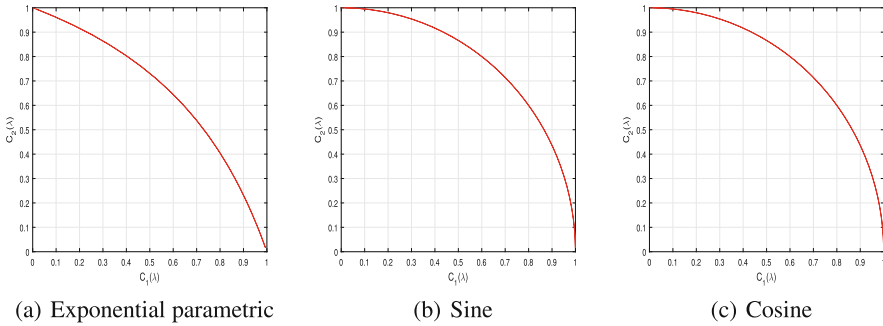


Fig. 2. Coefficients’ inverse relationship in the proposed mapping

3.1 Exponential Parametric Mapping

In Subsect. 2.3, a new parameter m allows to add one degree of freedom. Also, the exponential function can be extended by adding a new parameter a , this approach is called exponential parametric mapping. In this sense, coefficients’ expressions are $C_1^{exp_p}(\lambda) = \frac{e^{a\lambda^m}}{1-e}$ and $C_2^{exp_p}(\lambda) = \frac{e^{a(1-\lambda)^m}}{1-e}$, where $a \in \mathbb{R}$ and $m \in \mathbb{N}$. It can be readily proved that this approach also fulfills the homotopy conditions. Figure 2(a) shows the plotting of coefficients when varying the value of λ .

3.2 Sinusoidal Sine Mapping

The use of trigonometric functions is inspired by the wide range of sine and cosine function applications, including applications in astronomy and geodesy to determine inaccessible distances, its use in alternating electrical current and its applications for AM and FM modulation, among others. Coefficients’ expressions are $C_1^{sin}(\lambda) = \sin(\lambda^m \frac{\pi}{2})$ and $C_2^{sin}(\lambda) = \sin((1-\lambda)^m \frac{\pi}{2})$, where $m \in \mathbb{N}$. These coefficients fulfill the homotopy conditions and Fig. 2(b) shows its inverse relationship.

3.3 Sinusoidal Cosine Mapping

A cosine function is a $\frac{\pi}{2}$ offset sine function, therefore it is also susceptible to be used as a homotopy function. Accordingly, coefficients' expressions using a cosine function can be expressed as $C_1^{cos}(\lambda) = \cos((1 - \lambda)^m \pi)$ and $C_2^{cos}(\lambda) = \cos(\lambda^m \pi)$, where $m \in \mathbb{N}$. The Fig. 2(c) shows the inverse relationship between the coefficients, this fulfill the homotopy conditions.

3.4 Polynomial Coefficient Added

In addition, all couple of coefficient can be modified by C_1^n y C_2^n , where $n \in \mathbb{N}$.

4 Quality Measure

To quantify the studied approaches performance we apply the quality measurement proposed in [1] which is based on a relation of a homotopic reference approach and a homotopic approach to evaluate. Mainly, this quality measure makes it possible to determine the degree to which the homotopy evaluated differs from the reference homotopy. The following considerations are taken into account when selecting the limits:

- When $m = \infty$ and $n = 1$, the first critical homotopy (FCH) is obtained, it presents a sharp deformation without smooth transition between objects or functions.
- When $m = 1$ and $n = \infty$, the second critical homotopy is obtained (SCH), it presents the same deformation of SCH.

Therefore, the relationship is made between the reference homotopy (when $m = 1$ and $n = \infty$, known as SFC case) that can have as homotopic function any of those that were addressed in this study. As expected, if we take the case when $n = 1$ and $m = \infty$ we can determine the lower limit, also known as FCH case. Figure 3(a) shows the limits. From a geometric point of view, an inverse ratio curve of the coefficients must be within the aforementioned range of curves.

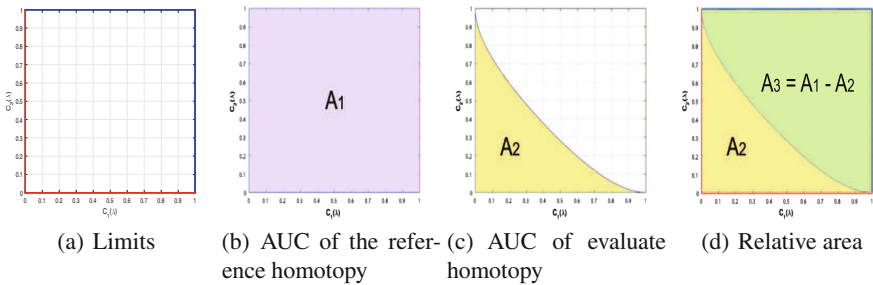


Fig. 3. Figure 3(a) shows the established reference limits. Upper reference (blue limit) with $n = \infty$ and $m = 1$, and lower reference (red limit) with $n = 1$ and $m = \infty$. Figure 3(b) to (d) show the area under curve (AUC) for quality measure. (Color figure online)

Consequently, the AUC between the reference homotopy (A_1) and the evaluation homotopy (A_2) are compared, generating an area of relation (A_3) between the two aforementioned homotopy approaches (see Fig. 3). Then, the quality measurement is given by $\eta = \frac{A_3}{A_1} = \frac{A_1 - A_2}{A_1} = \left(1 - \frac{A_2}{A_1}\right) \times 100\%$.

5 Experimental Setup

5.1 Data Sets and Methods

Experiments are carried out over four conventional data sets. Figures 4(a) and (b) show two artificial shells (spherical- and cylindrical-like, both with 1050 data points). As well, some images are considered, namely, six faces from the Yale Face database [9]. We pick pairs of faces to generate a set according to groups. The first group is represented by Fig. 4(c) and (d). The second group is represented by Fig. 4(e) and (f). The third group is represented by Fig. 4(g) and (h). We also noticed that given the existence of four faces we could generate $\mathbf{P}_2^4 = 12$ new groups including those already generated (first and second group).

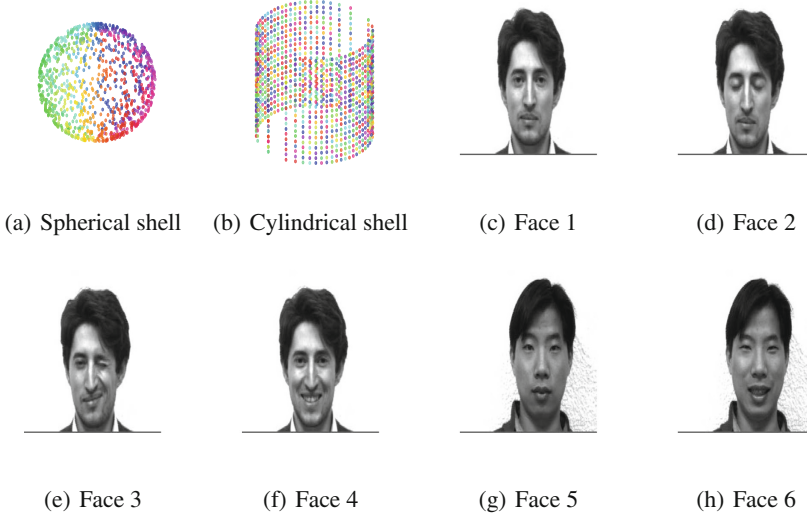


Fig. 4. Datasets. This figure depicts all datasets consider for experiments.

5.2 Methods to be Compared

We assess the homotopy mappings proposed in the Sect. 3, which are exponential parametric approach (ExparH), sinusoidal sine approach (SinH) and sinusoidal cosine approach (CosH). In addition, we assess the linear approach (LinH), polynomial approach (PolH), generalized exponential approach (GExpH) explained in Sect. 2. All of them are tested over all considered databases.

5.3 Parameter Settings

To compare the assessed homotopy mappings, we sweep the values of the homotopy parameter and calculate the mapping function h at each iteration. To do so, we define a vector $\Lambda = \{\lambda_i\}$ gathering the values of the homotopy parameter $\lambda_i \in \{0 : 0.02 : 1\}$, where notation $a_1 : \delta : a_2$ denotes a linearly increasing vector holding values ranged into the interval $[a_1, a_2]$ with a step size Λ . For every value of λ its corresponding ordered pair $(C_1(\lambda), C_2(\lambda))$ is calculated, which in turn is associated with an instantaneous mapping function h . In this case, this function represents an interpolated image (just as it was a frame of a video).

6 Results and Discussion

From Fig. 5, we can observe the behavior of the inverse relationship between coefficients C_1 and C_2 . These curves were made for different values of m (homotypic levels) for the same homotopic function. In each level of homotopy, its corresponding value can be computed.

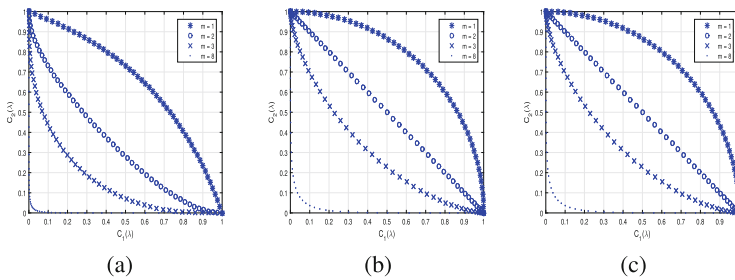


Fig. 5. $C_1(\lambda)$ vs $C_2(\lambda)$ plotting at different levels of m . Figure 5(a) for the proposed homotopic parametric exponential function of the Sect. 3. Figure 5(b) and (c) for the proposed homotopic sinusoidal-based functions of the Sect. 3.

In Fig. 5(a) it can be seen that the parametric exponential function presents a greater degree of freedom than the generalized exponential (Fig. 1(d)) proposed in figure [1], this is possible due to the inclusion of parameter a . Figures 5(b) and (c) depict different levels for the proposed functions, sine and cosine. It can be seen that for the same levels of m the obtained curves are the same so that the results before the same m without variations in n are equal. Our proposed model involves the inclusion of a parameter n . From Figs. 6, 7 and 8 illustrates the behaviour of coefficients for the three homotopic functions proposed in Sect. 3. Since the proposed parameterized exponential function has a parameter a , it is also varied and the results of the inverse relationship of the coefficients in Fig. 9 are presented. Results of applying the parameter-based exponential homotopic function in the deformation of the datasets are shown in the Fig. 10. For this, we consider $m = 1$, $a = -1$ and $n = 1$.

Figure 11 shows the results of applying the sinusoidal homotopic functions in the deformation of the datasets. We only present the results of the sine function because in both sine and cosine functions the results are the same under the consideration of $m = 1$ and $n = 1$. When we consider $m = 2$ and $n = 2$, the result is present in the Fig. 12(d) for all functions. Finally, when we consider $m = 4$ and $n = 2$, and $a = -3$ in addition for parametric exponential function, we obtain the results presented in Fig. 12(d). From the foregoing, it can be affirmed that the inclusion of parameters and new homotopic functions allow the generation of different curves between the coefficients and thus different deformations can be achieved.

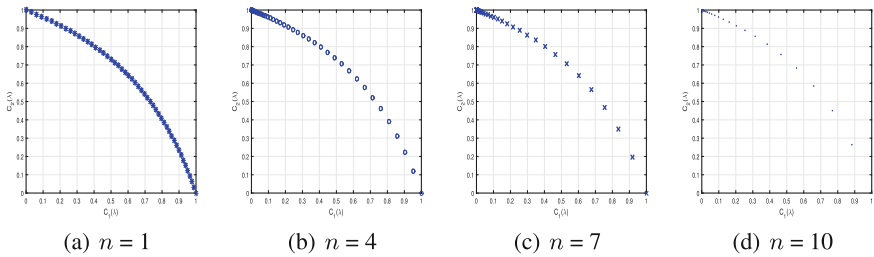


Fig. 6. Parameter n variation in the exponential parameter function, $m = 1$.

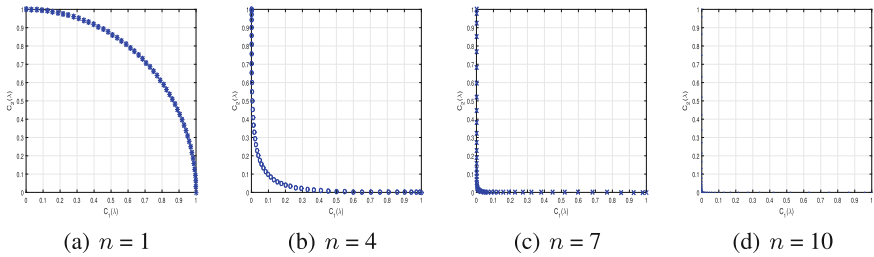


Fig. 7. Parameter n variation in the sine function, $m = 1$.

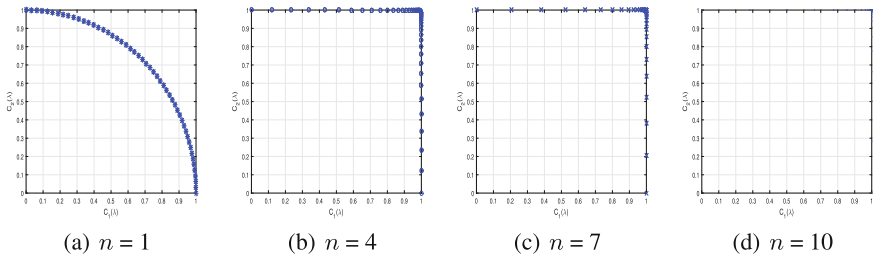


Fig. 8. Parameter n variation in the cosine function, $m = 1$.

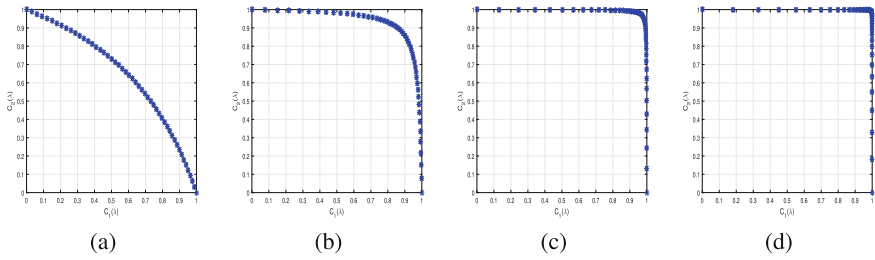


Fig. 9. Parameter a variation in the exponential parameter function, $m = 1$ and $n = 1$.

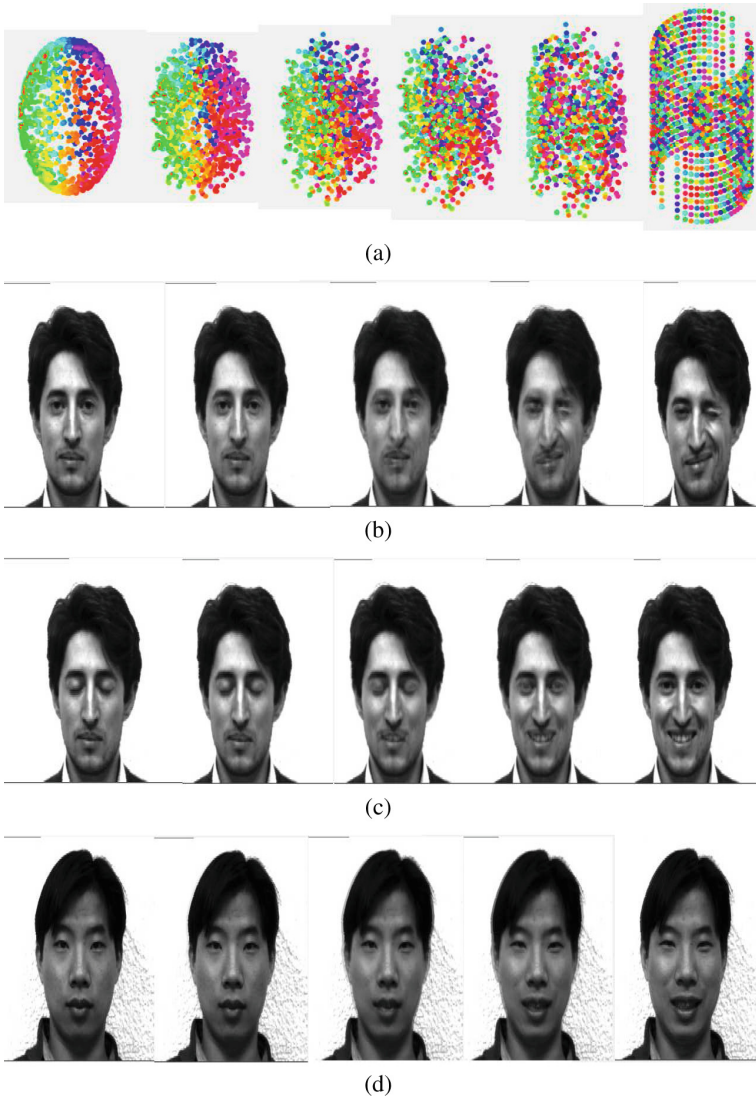


Fig. 10. Results of parametric exponential function with $m = 1$, $a = -1$ and $n = 1$.

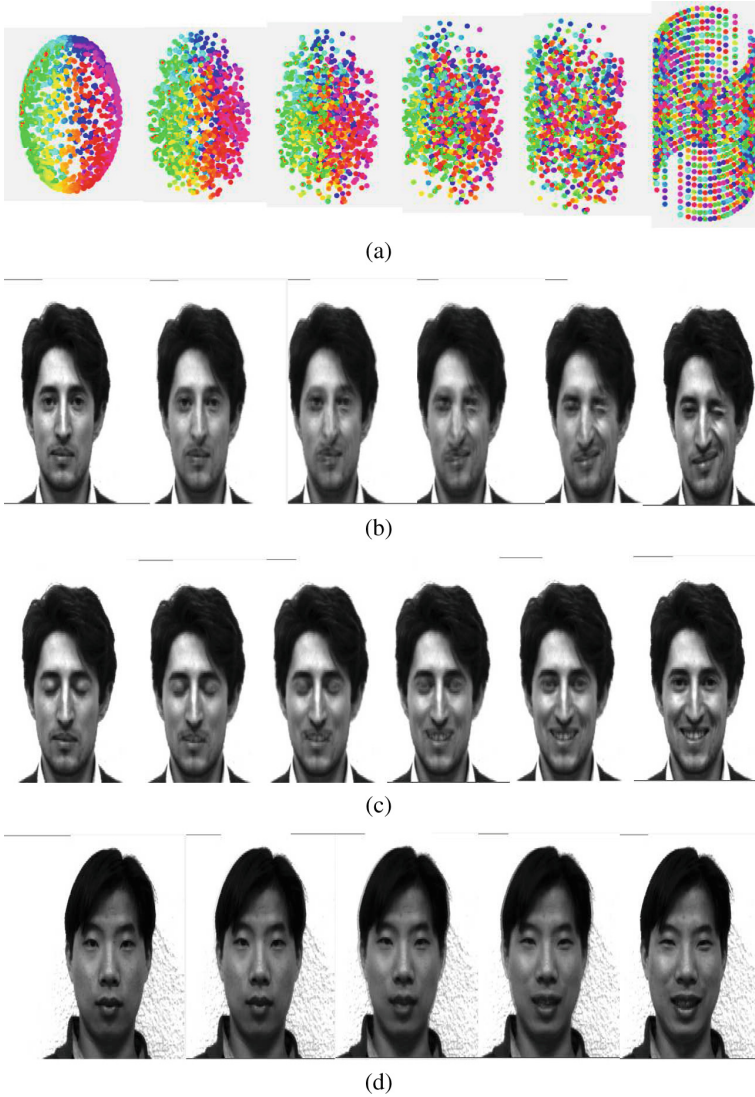


Fig. 11. Results of sine (and cosine, due to in this case the results are the same) function with $m = 1$ and $n = 1$.



(a)



(b)



(c)



(d)

Fig. 12. Results of all functions with $m = 2$ and $n = 2$. Figure 12(a) for sine function, Fig. 12(b) for cosine function, Fig. 12(c) for parametric exponential function with $a = -1$ and Fig. 12(d) for parametric exponential function with $m = 4$, $n = 2$ and $a = -2$

7 Final Remarks

The processes of deformation of objects by means of continuous transformation mapping from one to another can be carried out by means of homotopy principles. Since real-world dynamic behaviors are susceptible to be represented/approximated by homotopic functions, the expansion of coefficients

through the integration of new parameters allows for increasing the degrees of freedom that benefit the model by generating close-ups of dynamic behaviors. In this work, the exponential functions are parameterized. In addition, two new functions are introduced to represent the coefficients involved in the deformation, using sinusoidal functions. The homotopy that has been proposed here proves to be adequate by presenting deformations with slight variations. The quality measurement allows us to choose the parameters that generate a better deformation, so a smooth deformation was reached (particularly, by the cosine-based approach) when resultant η lies within the interval $[0.85, 0.92]$.

As a future work, on one hand, we are aiming to explore new homotopic mappings and parameterizations able to generate realistic deformations by including/tuning/optimizing free parameters. On the other hand, functions to deform more than two objects are to be further studied.


Acknowledgments. This work is supported by the “Smart Data Analysis Systems - SDAS” group (<http://sdas-group.com>).

References

1. Alpala, L.O., Peluffo-Ordóñez, D.H., González-Castaño, C., Guasmayan, F.A.: Deforming objects via exponential homotopy: a first approach. In: 2015 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA), pp. 1–6. IEEE (2015)
2. Riehl, E.: *Categorical Homotopy Theory*, vol. 24. Cambridge University Press, Cambridge (2014)
3. He, J.H.: Homotopy perturbation method: a new nonlinear analytical technique. *Appl. Math. Comput.* **135**(1), 73–79 (2003)
4. Babolian, E., Azizi, A., Saeidian, J.: Some notes on using the homotopy perturbation method for solving time-dependent differential equations. *Math. Comput. Modell.* **50**(1), 213–224 (2009)
5. Molina-Abril, H., Real, P.: A homological-based description of subdivided nD objects. In: Real, P., Diaz-Pernil, D., Molina-Abril, H., Berciano, A., Kropatsch, W. (eds.) CAIP 2011. LNCS, vol. 6854, pp. 42–50. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23672-3_6
6. Peluffo-Ordóñez, D.H., Alvarado-Pérez, J.C., Lee, J.A., Verleysen, M., et al.: Geometrical homotopy for data visualization. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2015) (2015)
7. Salazar-Castro, J., Rosas-Narváez, Y., Pantoja, A., Alvarado-Pérez, J.C., Peluffo-Ordóñez, D.H.: Interactive interface for efficient data visualization via a geometric approach. In: 2015 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA), pp. 1–6. IEEE (2015)
8. Feng, X.: *Geometric and topological modeling techniques for large and complex shapes*. Michigan State University (2014)
9. Georgiades, A., Belhumeur, P., Kriegman, D.: Yale face database, vol. 2. Center for computational Vision and Control at Yale University (1997). <http://cvc.yale.edu/projects/yalefaces/yalefa>



Thermal Imaging for Localization of Anterior Forearm Subcutaneous Veins

Orcan Alpar and Ondrej Krejcar^(✉) 

Faculty of Informatics and Management, Center for Basic and Applied Research,
University of Hradec Kralove, Rokitanskeho 62,
500 03 Hradec Kralove, Czech Republic
orcanaalpar@hotmail.com, ondrej.krejcar@uhk.cz

Abstract. The anterior forearm recognition systems are very popular identify the subcutaneous veins, mostly by devices operating real-time. Real time projection systems are carried out by handheld devices with a near infrared (NIR) camera and a laser projector to choose venipuncture sites; however what we propose in this paper is an easier and more reliable way using infrared thermal (IR-T) camera. At the forearm some veins like Cephalic vein are mostly so concealed to detect by NIR cameras, therefore we propose an alternative method for localization of the whole vein system without preprocessing. Briefly in the thermograms, the forearm is segmented from the surrounding by crisp 2-means and the vein system is reconstructed on blank images by directional curvature method. All directional curvatures are combined by addition for merging the layers and highlighting mutual veins. The preliminary results are promising since all the vein system is revealed with invisible veins without any preprocessing.

Keywords: Thermal imaging · Forearm veins · Directional curvature
Crisp 2- means · Localization · Segmentation

1 Introduction

Finding venipuncture sites for blood sampling and invasive medication could be rather easy for experts; while assistive systems could provide additional screening to locate superficial veins. Although the superficial veins on the forearm are mostly visible or felt by fingertips of the experts, the veins could be undetectable due to lack of contrast, or presence of tattoos or due to skin color and various external factors. In these cases, visible spectrum cameras could not provide necessary visualization, therefore the researchers used near infrared imaging (NIR) for instance [1, 2] for better identification. On the contrary, based on our previous experience, it is easy to state that deeper veins are completely invisible in NIR imaging and superficial veins surely need preprocessing to be revealed.

Given these facts, we propose a novel forearm vein identification system using thermograms capturing the radiation emittance which is comprehended as temperature. Above all, the thermograms are red-green-blue (RGB) images depending on the pseudocoloring and some pseudocolored thermograms have very high contrast between

the background and the radiation emitting tissues. In additions although the original thermograms seem grayscale, they consist of RGB layers when extracted. Therefore, the forearm on the images are firstly segmented by crisp 2-means by high contrast thermography. Subsequently, we turned the segmented images into complementary and the vein system is reconstructed on black frames by directional curvature method for fusion.

In the literature, there are several papers published in recent years on vein system estimation and reconstruction by thermal imaging: Buddharaju and Pavlidis [3] proposed superficial face veins estimation by thermal imaging and anisotropic diffusion. Extracting the radiation distances, they constructed the vein system on the face images successfully including the invisible veins. In our previous research paper, [4] we initiated this research by proposing a methodology for vein estimation of dorsal hands.

Briefly, Huang et al. [5] used NIR imaging to identify vein system of dorsal hands and proposed fusion of global local texture and shape based patterns. Joarder et al. [6] dealt with grayscale thermograms of dorsal hands and estimated vein patterns by median filtering and artificial bee colony algorithm. Yun-Peng et al. [7] focused on dorsal hands for identification of the vein system by dynamic thresholding and approximated the veins by binarization and thinning. Yan et al. [8] clustered the hands and tried to identify the veins inside of the cluster windows by SIFT algorithm. In addition, there are many papers on vein detection by several imaging methods, also including hand and finger veins, such as:

Lee et al. [9] focused on hand vein biometrics and detected dorsal hand veins by NIR imaging and two dimensional Gabor filters. Sheetal and Goela [10] dealt with NIR based dorsal fist vein recognition by filtering and contrast enhancements, even though the veins are totally visible. Again with biometric purposes, Janes and Júnior [11] prepared a hand stabilizer with a set of LED to enhance NIR images of the hands. However, they also had to apply contrast enhancement before finding region of interest.

Moreover, Sekthivel [12] investigated the proper wavelengths for NIR imaging to detect palm veins by a specific device they constructed for biometric purposes. They applied median filter for enhancing the images; although the vein system in the NIR images they extracted are one of the most visible system despite very low resolution. There are more papers on dorsal hand and palm vein recognition [13–16] and finger vein detection [17]. On the other hand, some references are very helpful for image enhancement techniques in vein recognition such as: [18–22]. In addition, Buddharaju et al. published two more papers on Physiology based face recognition using thermal imaging which are all worth mentioning [23, 24].

Among these paper mentioned above, our kernel is totally different, since by they finger veins or dorsal hand veins, all of them are totally visible even with visible spectrum cameras. Our basic difference is to reveal the invisible veins in the forearm region; not possible to find in fingers or hands.

The paper starts with preliminaries of forearm anatomy and NIR imaging vs. thermography with various pseudocolored thermogram of forearm. Moreover, the segmentation process from the image by crisp 2 means and several experiments based on forearm thermography are presented; and finally the paper ends with conclusion and discussion.

2 Forearm Thermography

In this section, we firstly present the vein anatomy of forearms, even though the vein patterns actually are biometric characteristics which all are unique. Afterwards, we will present our NIR images and thermograms to show the basic differences.

2.1 Forearm Anatomy

The veins in forearms run from the elbow trough the wrist and fingers carrying blood to and from the extremities of the arm. There are two major veins in the forearm vein system: Cephalic and Basilic veins which interconnects with median Cubital vein at the elbow. Although it is mostly visible at upper arm, accessory Cephalic and main Cephalic veins become invisible; despite the Basilic vein while passing through forearms. The basic anatomy of the vein system ss presented in Fig. 1.

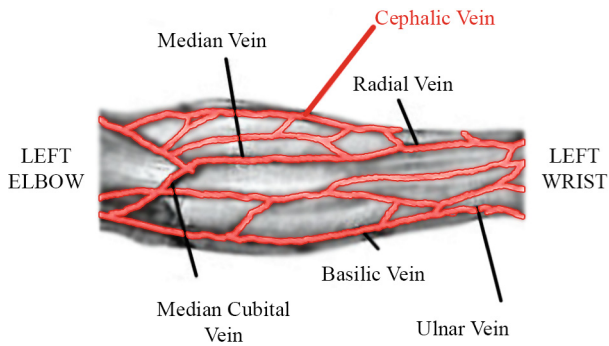


Fig. 1. Forearm vein anatomy

As mentioned above, all the vein patterns are unique and represent the characteristics trait of an individual like finger prints and hand veins. However, the veins presented in Fig. 1 more or less should be at very similar locations; even if the patterns are totally different.

2.2 Forearm NIR Imaging

NIR imaging is a promising methodology to reveal the vein system; yet it has a major drawback on Cephalic vein recognition. At the forearm, Cephalic vein mostly loses its visibility and neither visible spectrum nor NIR cameras could identify the Cephalic vein; unless the focus is at the vein, which could be seen in Fig. 2.

Therefore as an alternative approach, we suggest to use thermal imaging instead, when identification of the whole vein system is necessary. The main difference would be the radiation differences that the thermograms would give better results for the veins which are not so visible by NIR cameras. On the other hand, it is obviously not possible to apply any kind of vein recognition methodology without preprocessing of NIR imaging; while in IR-T, it is not necessary.



Fig. 2. Experiments with NIR camera

2.3 Forearm IR-T Imaging

All objects, organic or inorganic, emit heat radiation in infrared region of the spectrum that can be captured by thermal cameras since the visible spectrum cameras are not able to recognize this radiation. Therefore thermography is a very common diagnostic tool to identify contrast in radiation levels. In clinical medicine or biomedicine, like in our research, it is also very promising methodology for diagnosis of cancer, inflammation, nerve entrapments and such.

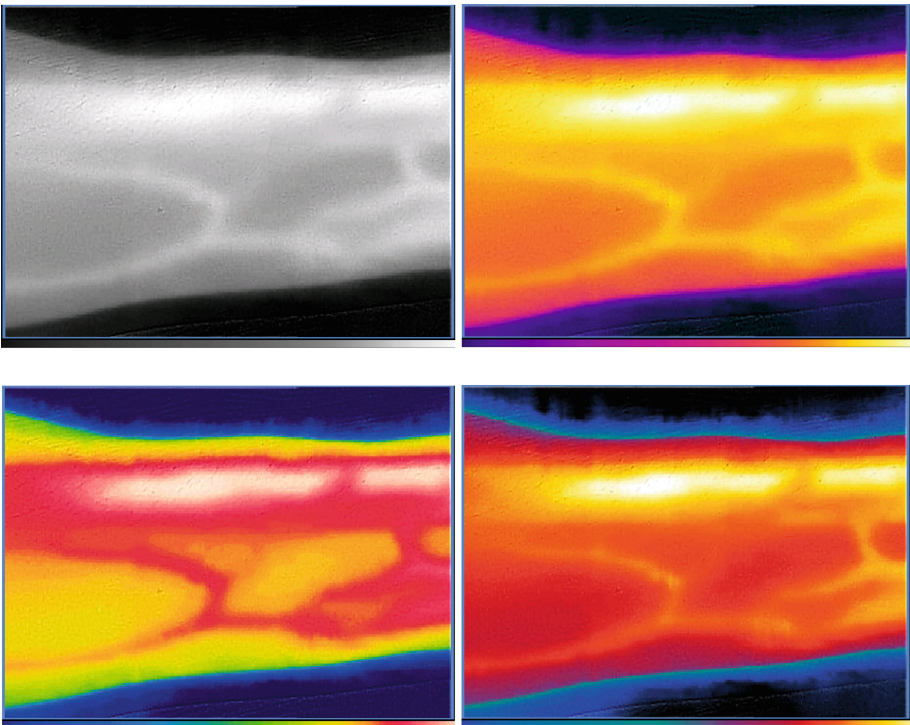


Fig. 3. Pseudocolored thermograms (original on top-left, iron on top-right, rainbow on bottom-left, lava on bottom-right) (Color figure online)

Depending on the specifications of the thermal camera, it is possible to capture only the radiation neglecting photons, therefore identification of thermal differences is possible in any light conditions. While capturing, adjusting the contrast is still plausible when working on a specific contrast pattern. For instance leakage on a wall would be very simply recognized by setting a low radiation thermogram; or an inflammation would be revealed better with a high radiation pseudocoloring. The common pseudocolored thermograms of FLIR infrared thermal camera with 0.96 emissivity, from 20 cm distance are presented in Fig. 3.

These thermograms could give better results depending on the context mentioned above, we will deal with original thermograms which seems to be grayscale; however they actually consist of three layers, red, green and blue when extracted. Any pixel $p_{i,j,k}$ on an image of original thermogram $R_{i,j,k}$ could be stated as:

$$p_{i,j,k} \in R_{i,j,k} \quad (i = [1 : w], j = [1 : h], k = [1 : 3]) \tag{1}$$

where w is the width, j is the height, k is the color layer. Above all, we need to segment the forearm from the surrounding for better identification of the vein system.

3 Segmentation

Since there is huge contrast between the forearm and its surroundings, the common and feasible way for segmentation is crisp 2 means. When clustering an RGB image into two binary sections, the objective function is

$$J = \sum_{a=1}^2 \sum_{b=1}^{i,j,k} u_{a,b}^m d_{a,b} \tag{2}$$

where $u_{a,b}$ are memberships, i, j, k are the data points in $R_{i,j,k}$ and $d_{a,b}$ is the squared Euclidean distance between data points in X_b matrix and C_b clusters. In this case, $u_{a,b}$ should satisfy:

$$\sum_{b=1}^2 u_{a,b} = 1 \tag{3}$$

While $d_{a,b} > \eta$, $u_{a,b}$ is represented as:

$$u_{a,b} = \left((d_{a,b})^{1/(m-1)} \sum_{s=1}^2 \left(\frac{1}{d_{a,s}} \right)^{1/(m-1)} \right)^{-1} \tag{4}$$

The average values of two clusters could be found by:

$$[C_b] = \sum_{a=1}^N u_{a,b}^m X_b / \sum_{a=1}^N u_{a,b}^m \tag{5}$$

which is also a palette matrix $[C_b]$. This palette matrix gives the results of segmentation as the average values of each cluster, namely:

$$[C_b] = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \tag{6}$$

however we need to change the color codes into binary by,

$$[\dot{C}_b] = \begin{cases} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & a+b+c > d+e+f \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} & a+b+c < d+e+f \end{cases}, \tag{7}$$

The final segmented image $E_{i,j}$ could easily be found by dot product $[\dot{C}_b]$ with of the image matrix, presented in Fig. 4.

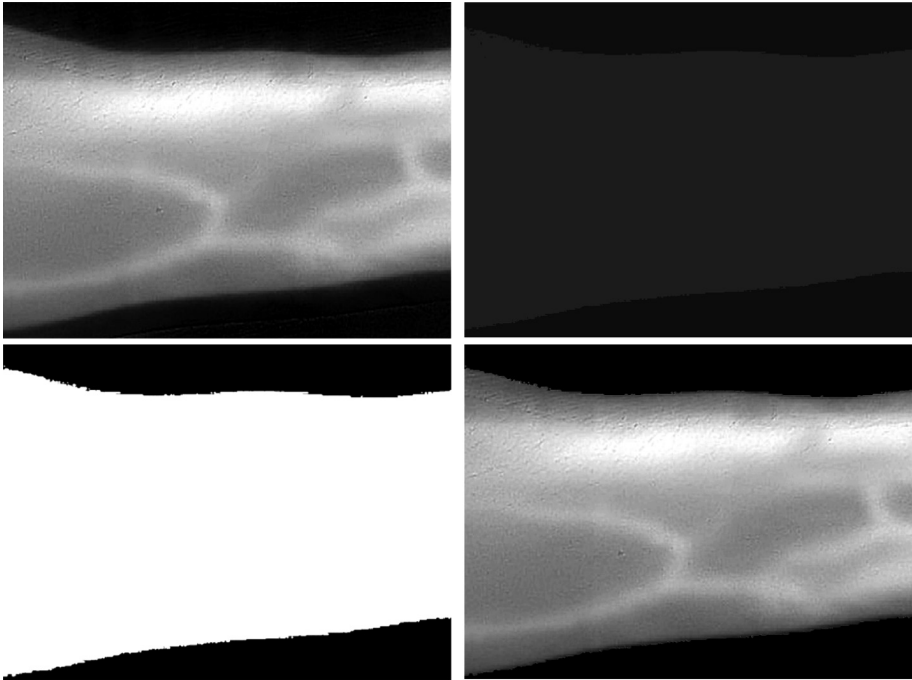


Fig. 4. Segmentation process (original image $R_{i,j,k}$ on top-left, image with $[C_b]$ on top-right, image with $[\dot{C}_b]$ on bottom-left, segmented image $E_{i,j}$ on bottom-right) (Color figure online)

Another way for segmentation is red channel conversion of pseudocolored thermograms we used in [4], which is rather easy process that the red channel carries the information of high radiation regions.

4 Directional Curvature Method

The maximum curvature method starts with calculating the curvatures to find the summations in each direction, horizontal, vertical and two diagonal directions. Afterwards, in each direction maximum curvatures are computed to connect the vein centers. We only found the directional curvatures in this paper and omitted the maximum curvatures and combining them. Given that $E_{i,j}$ is the segmented image and $E_{i,j}(x, y)$ is the pixel intensity, the cross sectional profile $P_f(z)$ is gathered from $E_{i,j}(x, y)$ in four directions and positions where z represents the position. To find the relation of $P_f(z)$ vs $E_{i,j}(x, y)$, a mapping function M is determined as $E_{i,j}(x, y) = M(P_f(z'_i))$. The curvatures are found by the second order derivatives by:

$$\kappa(z) = \frac{\frac{d^2 P_f(z)}{dz^2}}{\left(1 + \left(\frac{dP_f(z)}{dz}\right)^2\right)^{\frac{3}{2}}} \quad (8)$$

If $\kappa(z) > 0$ it is supposed to be concave, therefore the positions of the points are saved as z'_i where $i = 0, 1, \dots, N - 1$ and N is the number of local maximum points. Subsequently the scores are assigned to the center positions as follows:

$$S(z'_i) = \kappa(z'_i) W_r(i) \quad (9)$$

where $W_r(i)$ is the width of the region. Each score obtained by this procedure is assigned to a blank image V as follows:

$$V(x'_i, y'_i) = V(x_i, y_i) + S(z'_i) \quad (10)$$

where x'_i, y'_i is the points defined by $E_{i,j}(x', y') = M(P_f(z'_i))$. In maximum curvature, the operation applied to all pixels is defined as:

$$C_{d1}(x, y) = \min\{\max(V(x+1, y), V(x+2, y), V(x-1, y), V(x-2, y))\} \quad (13)$$

Also the other three directions $C_{d2}(x, y), C_{d3}(x, y), C_{d4}(x, y)$ could be found in similar way. The new image could be reconstructed by $G = \max(C_{d1}, C_{d2}, C_{d3}, C_{d4})$ in traditional maximum curvature approach; however, we only reconstructed the blank images in each direction and the summations are analyzed, which could be seen in Fig. 5.

These images represent the directional curvatures our algorithm managed to reconstruct on blank images. Form the images, the pattern could be easily detected that each image consists of one direction.

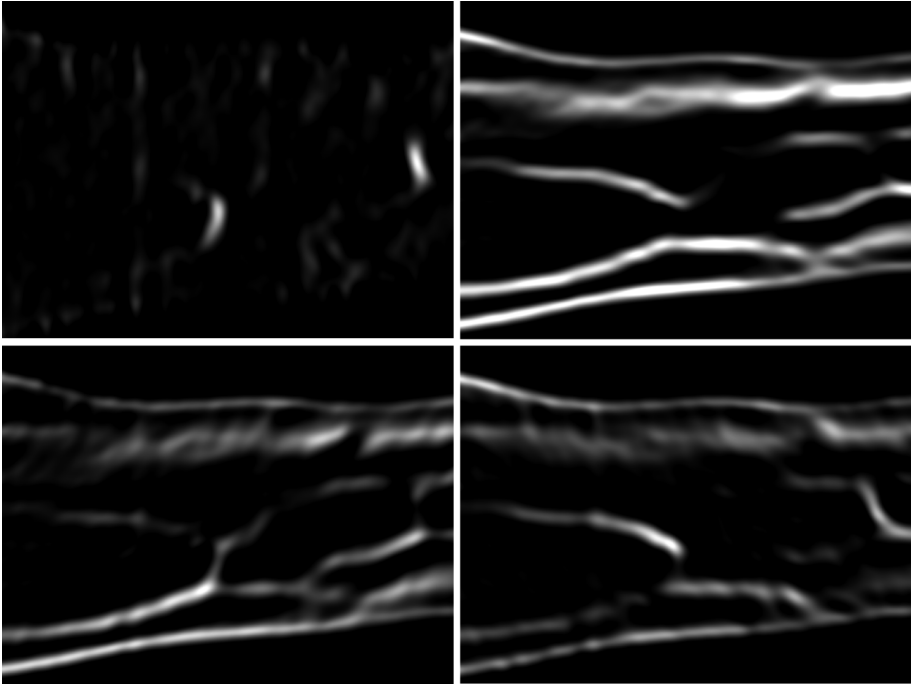


Fig. 5. $C_{d1}(x, y)$ on top-left, $C_{d2}(x, y)$ on top-right, $C_{d3}(x, y)$ on bottom-left, $C_{d4}(x, y)$ on bottom-right

5 Experimental Results

Depending on the directions of thermal identification device, we determined four types of combination of the operations. Considering that $C_{d1}(x, y)$ is vertical, $C_{d2}(x, y)$ is horizontal and $C_{d3}(x, y)$, $C_{d4}(x, y)$ are the diagonal curvatures the plausible combinations are $C_{d1}(x, y) + C_{d2}(x, y)$, $C_{d2}(x, y) + C_{d3}(x, y)$, $C_{d3}(x, y) + C_{d4}(x, y)$ and $C_{d2}(x, y) + C_{d3}(x, y) + C_{d4}(x, y)$ for forearm vein system. The results are presented in Fig. 6.

According to the experiment and depending on the combinations, it is seen that the detail level and the connections are changing. $C_{d2}(x, y)$, representing the horizontal reconstruction, is totally clear while the whole vein system couldn't be constructed due to diagonal veins. The other images gathered by the summations mentioned above are more close to reality, therefore as long as we achieve the layers C_{d*} , any combination of them including summation or averaging seemingly would work.

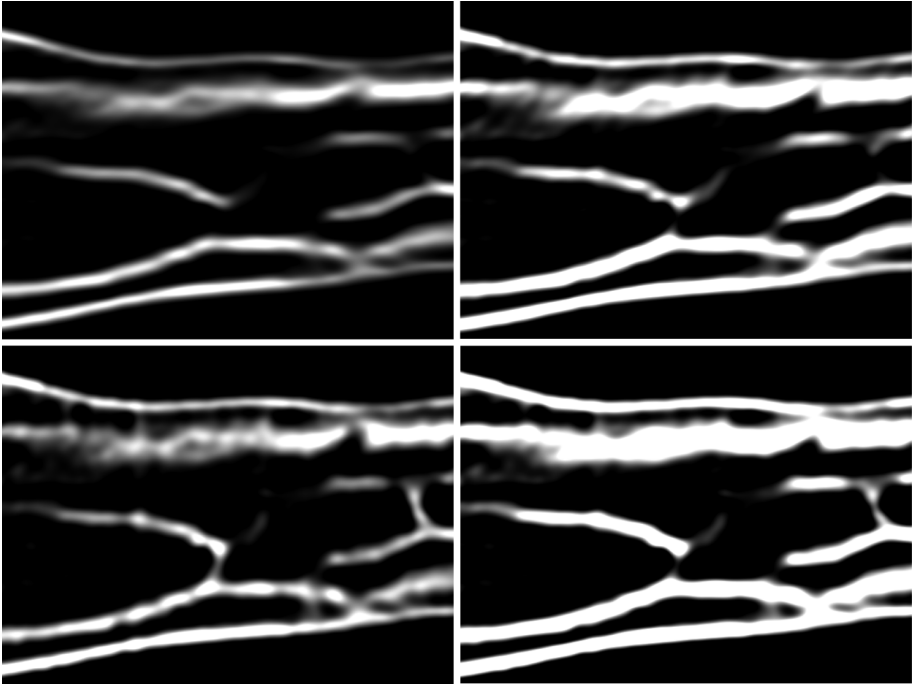


Fig. 6. $C_{d1}(x, y) + C_{d2}(x, y)$ on top-left, $C_{d2}(x, y) + C_{d3}(x, y)$ on top-right, $C_{d3}(x, y) + C_{d4}(x, y)$ on bottom-left, $C_{d2}(x, y) + C_{d3}(x, y) + C_{d4}(x, y)$ on bottom-right

6 Evaluation of the Results

In this section we present the differences between the maximum curvature and our method by comparing the results of each algorithm on the same image. Aside from our papers, there is only one paper deals with estimating vein system by maximum curvature [3]; so our results could be compared with this book chapter. They used maximum curvature mainly for face recognition; yet in a section of this book chapter, they used a similar maximum curvature algorithm. Although this method is very useful for detecting the paths so the vein patterns; it however doesn't give any information about the thickness of the veins; which could be seen in Fig. 7.

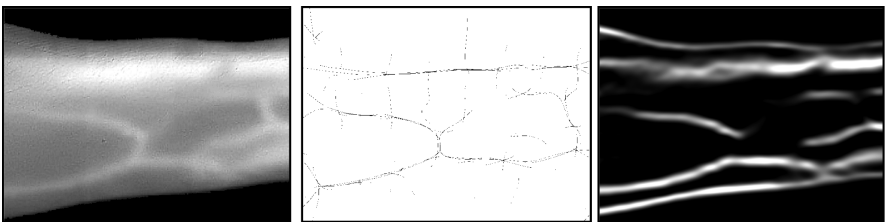


Fig. 7. Segmented image $E_{i,j}$ on the left, maximum curvature G in the middle, $C_{d1}(x, y) + C_{d2}(x, y)$ on the right.

Any kind of benchmarking between IR-T and NIR imaging is clearly not possible, since NIR camera couldn't even recognize the Cephalic vein.

7 Conclusions and Discussions

Despite the benefits of the alternative approach we proposed, we clearly do not state that this method is totally superior ever; yet it is so promising for detecting the veins using radiation differences that NIR cameras cannot. Therefore our first concern is finding alternatives when NIR imaging cannot fulfill the requirements. Finding directional curvatures with merging is an addition to this concern to finalize this preliminary research

There is no major weakness of the procedure we presented; yet we had some assumptions about resolution of the thermal camera, distance and temperature conditions. The images extracted from the thermal camera are in 1530×2040 in 24° room temperature from a 20 cm distance. However the main assumption is the contrast that veins carrying blood should be warmer than the skin tissues as usual. Therefore the main requirement is the temperature of the place where this kind of experiment would be conducted.

The outcomes of this research could be applied as a handheld device that will recognize the invisible veins by thermal camera and project to the patient's forearm. In this case, if the device is moving from left to right, $C_{d2}(x, y)$ would be sufficient; while if from bottom to top, $C_{d1}(x, y)$ could be applicable. Therefore, the results and the methodology mentioned in this paper could shed light on an adaptive device that will recognize the subcutaneous veins better than other imaging techniques. If necessary, some other algorithms could be used for segmentation [25], that we had presented in our previous paper.

Acknowledgment. The work and the contribution were supported by the SPEV project "Smart Solutions in Ubiquitous Computing Environments", University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic. We are also grateful for the support of Ph.D. students of our team (Ayca Kirimtat and Pavel Blazek) in consultations regarding application aspects.

References


1. Kim, D., Kim, Y., Yoon, S., Lee, D.: Preliminary study for designing a novel vein-visualizing device. *Sensors* **17**(2), 304–323 (2017)
2. Song, J.H., Kim, C., Yoo, Y.: Vein visualization using a smart phone with multispectral Wiener estimation for point-of-care applications. *IEEE J. Biomed. Health Inform.* **19**(2), 773–778 (2015)
3. Buddharaju, P., Pavlidis, I.T.: Physiology-based face recognition in the thermal infrared spectrum. In: *Medical Infrared Imaging: Principles and Practices*, pp. 18.1–18.16. CRC press (2012)

4. Alpar, O., Krejcar, O.: Superficial Dorsal hand vein estimation. In: Rojas, I., Ortuño, F. (eds.) IWBBIO 2017 Part I. LNCS, vol. 10208, pp. 408–418. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56148-6_36
5. Huang, D., Zhang, R., Yin, Y., Wang, Y.: Local feature approach to dorsal hand vein recognition by centroid-based circular key-point grid and fine-grained matching. *Image Vis. Comput.* (2016). <http://dx.doi.org/10.1016/j.imavis.2016.07.001>
6. Joardar, S., Chatterjee, A., Bandyopadhyay, S., Maulik, U.: Multi-size patch based collaborative representation for Palm Dorsa Vein Pattern recognition by enhanced ensemble learning with modified interactive artificial bee colony algorithm. *Eng. Appl. Artif. Intell.* **60**, 151–163 (2017)
7. Yun-peng, H., Zhi-yong, W., Xiao-ping, Y., Yu-ming, X.: Hand vein recognition based on the connection lines of reference point and feature point. *Infrared Phys. Technol.* **62**, 110–114 (2014)
8. Yan, X., Kang, W., Deng, F., Wu, Q.: Palm vein recognition based on multi-sampling and feature-level fusion. *Neurocomputing* **151**, 798–807 (2015)
9. Lee, J.C., Lee, C.H., Hsu, C.B., Kuei, P.Y., Chang, K.C.: Dorsal hand vein recognition based on 2D Gabor filters. *Imaging Sci. J.* **62**(3), 127–138 (2014)
10. Sheetal, R.P., Goela, K.G.: Image processing in hand vein pattern recognition system. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **4**(6), 427–430 (2014)
11. Janes, R., Júnior, A.F.B.: A low cost system for Dorsal hand vein patterns recognition using curvelets. In: *Proceedings of the 2014 First International Conference on Systems Informatics, Modelling and Simulation*. IEEE Computer Society, pp. 47–52 (2014)
12. Sakthivel, G.: Hand vein detection using infrared light for web based account. *Int. J. Comput. Appl.* **112**(10), 17–21 (2015)
13. Premalatha, K., Natarajan, A.M.: Hand vein pattern recognition using natural image statistics. *Def. Sci. J.* **65**(2), 150–158 (2015)
14. Srivastava, S., Bhardwaj, S., Bhargava, S.: Fusion of palm-phalanges print with palmprint and dorsal hand vein. *Appl. Soft Comput.* **47**, 12–20 (2016)
15. Lee, J.C., Lo, T.M., Chang, C.P.: Dorsal hand vein recognition based on directional filter bank. *SIViP* **10**(1), 145–152 (2016)
16. Morales-Montiel, I.I., Olvera-López, J.A., Martín-Ortíz, M., Orozco-Guillén, E.E.: Hand vein infrared image segmentation for biometric recognition. *Res. Comput. Sci.* **80**, 55–66 (2014)
17. Miura, N., Nagasaka, A., Miyatake, T.: Extraction of finger-vein patterns using maximum curvature points in image profiles. *IEEE Trans. Inf. Syst.* **90**(8), 1185–1194 (2007)
18. Djerouni, A., Hamada, H., Loukil, A., Berrached, N.: Dorsal hand vein image contrast enhancement techniques. *Int. J. Comput. Sci.* **11**(1), 137–142 (2014)
19. Dan, G., Guo, Z., Ding, H., Zhou, Y.: Enhancement of Dorsal hand vein image with a low-cost binocular vein viewer system. *J. Med. Imaging Health Inform.* **5**(2), 359–365 (2015)
20. Wang, Y., Duan, Q., Shark, L.K., Huang, D.: Improving hand vein recognition by score weighted fusion of wavelet-domain multi-radius local binary patterns. *Int. J. Comput. Appl. Technol.* **54**(3), 151–160 (2016)
21. Wang, G., Wang, J., Li, M., Zheng, Y., Wang, K.: Hand vein Image Enhancement Based on Multi-Scale Top-Hat Transform. *Cybern. Inf. Technol.* **16**(2), 125–134 (2016)
22. Wang, J., Wang, G., Li, M., Du, W., Yu, W.: Hand vein images enhancement based on local gray-level Information histogram. *Int. J. Bioautom.* **19**(2), 245–258 (2015)
23. Buddharaju, P., Pavlidis, I.T., Tsiamyrtzis, P., Bazakos, M.: Physiology-based face recognition in the thermal infrared spectrum. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 613–626 (2007)

24. Buddharaju, P., Pavlidis, I.T., Tsiamyrtzis, P.: Physiology-based face recognition. In: IEEE Conference on Advanced Video and Signal Based Surveillance, AVSS 2005, pp. 354–359. IEEE (2005)
25. Alpar, O., Krejcar, O.: Dorsal hand recognition through adaptive YCbCr imaging technique. In: Nguyen, N.-T., Manolopoulos, Y., Iliadis, L., Trawiński, B. (eds.) ICCCI 2016 Part II. LNCS (LNAI), vol. 9876, pp. 262–270. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45246-3_25



Detection of Irregular Thermoregulation in Hand Thermography by Fuzzy C-Means

Orcan Alpar and Ondrej Krejcar^(✉) 

Faculty of Informatics and Management, Center for Basic and Applied Research,
University of Hradec Kralove,
Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic
orcanalpar@hotmail.com, ondrej.krejcar@uhk.cz

Abstract. From numbness to acral necrosis, vasospasms confining to the fingers are stimulating a wide variety of symptoms. Induced by several reasons, the constriction of small arteries is a severe disorder and should be diagnosed as early as possible. The major indicator of a mild vasospasm is reduced blood flow creating a nonhomogeneous radiation patterns in hand thermography before a total discoloration of the fingers. Therefore, in this paper, an identification methodology of the vasospasms is discussed, through the fuzzy c-means. Since the pseudocolored thermograms give crucial information of radiation variety, we segmented the hands from the images by hard k-means for subsequent analysis of low-radiated regions determined by fuzzy c-means. The isolated low-radiated regions in red-channel converted images also reveals ischemia as well as thermoregulation even in very milder cases.

Keywords: Thermoregulation · Vasospasm · Hand thermography
Vasoconstriction · Fuzzy C-means · Hard k-means

1 Introduction

A vasospasm affecting the hands mainly refers to general narrowing of a blood vessel due to contraction of the surrounding muscles. The symptoms may occur as a result of exposure to cold; underlying diseases or existence of consistent pressure or overwhelming stress. In milder form of vasospasms, the patient usually suffers from numb and cold hands due to decreased blood flow through fingers. In more severe cases, however, there would be a visible triphasic or biphasic color change in the fingers during the attack. In the final phase, a kind of visible damage could occur by developing acral necrosis or ulcers or gangrene in the affected area, which all could be irreversible.

There are several diseases and disorders related to vasospasms: Acrocyanosis [1], Erythromelalgia [2], Acroparesthesia [3] or Hypothenar Hammer Syndrome [4] while the reason of vasospasms is connected to Raynaud's phenomenon [5] in the absence of any trauma or underlying disorder. Acrocyanosis is a type of benign vasoconstriction bringing color change in hands intensified by emotional stress or exposure to cold. Increasing fever and pain, Erythromelalgia is an episodic disorder that affects the hands by vasodilatation and erythema. Acroparesthesia causes numbness in the fingers due to

vasospasm mostly triggered by an unknown source or nerve compression. Hypothenar eminence corresponds the muscles controlling the movement of little fingers therefore hypothenar hammer syndrome refers to the numbness due to vascular occlusion in this region.

However, impairment of the circulation is not always due to diagnosable diseases or disorders or triggered by a trauma. Without any underlying conditions, the vasospasm of arteries in hands refers to primary Raynaud's phenomenon; while triggered by a medical condition, it refers to secondary Raynaud's phenomenon. On the other hand, nerve entrapments in Guyon's canal or Carpal's tunnel may cause vasoconstriction as well [6]. In any case, early diagnosis is possible by infrared thermography designed for revealing radiation differences after a special thermoregulation experiment [7–12].

Thermoregulation experiment is based on the consecutive processes of cooling the hands and rewarming to stimulate the vasoconstriction and the vasodilatation phases. Therefore the most common method is the infrared thermography to identify abnormality in blood flow. Infrared thermography is a methodology to detect and monitor the radiation from energy emitting matters and tissues. The wavelength of the energy is too long to be detected by naked eye, yet the energy itself is comprehended as heat. Furthermore as the temperature of a tissue increases, the infrared radiation emitted by the tissue increases as well.

The main output of thermography is a grayscale image produced by an infrared thermal camera. Depending on the requirements, the image could be pseudocolored by a built-in colormap to reveal the radiation differences in visible spectrum by a temperature legend. It actually is very practical to handle the pseudocolored thermograms rather than grayscale images when dealing with the identification of vasospasms. There are two very basic reasons for using pseudocolor thermograms instead of grayscale images; contrast and availability. The grayscale thermogram is usually colored by a built-in colormap which enables high contrast as well as better automatic segmentation possibilities. Despite the 256-bit interval of grayscale, pseudocolor thermograms bring out RGB images consisting of red, green and blue channels which triples the existing information. Second advantage is availability since most of the libraries and books only include the pseudocolor thermograms which could be seen in recently published clinical medicine books and books chapters such as [13, 14].

On the other hand, some colormaps reveal the required features that the others are not even close to fulfill the requirements. For instance, lava thermograms are perfect for vein detection while original grayscale or the rainbow thermograms don't give any information about this. However, the rainbow thermograms are very useful for differentiation of the object from the environment by a very high contrast. Finally, most of the researchers seemingly tend to deal with pseudocolor thermograms instead of grayscale as mentioned in previous paragraphs.

We therefore propose a methodology for detection and for early diagnosis of vasospasms and vasoconstrictions by thermal image processing techniques including fuzzy C-means for identification of colder regions after the thermoregulation experiment. Starting by segmentation of the dorsal hands from Rainbow pseudocolored thermograms by fuzzy k-means, the segmented image is preprocessed and binarized to refer to the region of interest in Lava pseudocolored thermogram. Omitting the other channels, red channel is analyzed to reveal low-radiated regions by fuzzy C-means. The

output of this process is turned into vertical summation to identify whether the distribution of radiation after the thermoregulation experiment is homogenous or not.

Unfortunately, the utilization of fuzzy methodology in thermal imaging is limited by any means, yet the results are so promising. Snekhalatha et al. [15] investigated rheumatoid arthritis in hand thermography and applied fuzzy C-means for segmentation of the thermal images. They also proposed the Expectation-Maximization algorithm to reach super-imposed images and to compare with fuzzy method. The major difference is the region of interest, since they dealt with the mid-region of the palms of healthy subjects and patients, including the wrists. In 2014, Mohiyuddin et al. [16], published a very similar paper, however including the thermograms of the necks and faces of the subjects to identify rheumatoid arthritis by fuzzy C-means. Initiation of each research above is converting the built-in RGB (red-green-blue) thermograms into HSV color space (hue-saturation-value), which we didn't need to. As a prospective and prominent approach, Antoni and Ravi [17] presented a fuzzy inference system that detects preliminary cancer using adaptive neurofuzzy and neural networks as well. We couldn't come across any paper dealing with vasospasms thermal imaging and fuzzy clustering in the literature, therefore it could be easily stated that we put forward a new area of utilization of fuzzy in medicine and thermal imaging.

The paper starts with explanation of preparation phase including the experimental design of thermoregulation, preprocessing of thermograms, segmentation and finalization. Afterwards we presented the algorithm for detecting the low-radiated regions in red-channel of segmented hands and a supportive distribution analysis before results and conclusion sections.

2 Preparation

In this section, we describe the preparation phase for the main experiment in detail as well as preprocessing of thermograms for better comprehension of the protocol.

2.1 Design of Thermoregulation Experiment

The special experiments designed for identifying thermoregulation in hands, especially the durations and the temperatures, may vary indeed, yet the common steps are immersion in water and exposure to water and taking infrared thermal images of the dorsal hands after the warming period. Governed by hypothalamus, thermoregulation is a process including the stimulation of vasoconstriction in exposure to cold and vasodilatation in exposure to hot, for keeping the inner body temperature stable, in homeostasis. Therefore, cold induced thermoreceptors immediately trigger peripheral vasoconstriction by narrowing down the vessels. Shortly after cold water immersion when the hands are taken out of the water, the vessels start to enlarge increasing the blood flow, preferably the hands are in room temperature and in standing position. The infrared thermograms should be recorded sometime after this process to reach homeostasis.

As mentioned above, the parameters for this experiment is various among the papers published in recent years. Table 1 provides some protocols applied by the researchers for testing Raynaud’s phenomenon.

Table 1. Protocols

Authors	Parameters		
	Cooling time	Cooling degree	Warming time
Ring and Ammer [18]	1 min	20 °C	10 min
Dinsdale and Herrick [19]	N/A	23 °C	10 min
Ismail et al. [20]	2 min	10 °C	20 min
Foerster et al. [21]	1 min	16 °C	N/A
<i>This paper</i>	<i>1 min</i>	<i>10 °C</i>	<i>20 min</i>

Although the main purpose and applied procedures are very similar, some differences in the detailed protocols might be practical, depending on the research. While Ring and Ammer [18] have done the experiment just like we did, Dinsdale and Herrick [19] however warmed the hands of the subjects directly from room temperature 23 °C to a degree of 30 °C, instead of cold immersion by water. Moreover, Ismail et al. [20] performed the experiments with the subjects wearing latex gloves; while Foerster et al. [21] continuously recorded the fingertip temperature for analyzing the whole warming process, instead of waiting for the homeostasis.

Given the parameters, we preferred the lowest cooling time and temperature with highest waiting time among these researches as in Table 1. Moreover, no other supplementary equipment is used for thermoregulation experiment and the room temperature is calculated as 25 °C during the experiment. Throughout the experiments, all images we recorded and analyzed are in 1530×2040 resolution, captured by FLIR infrared thermal camera with 0.96 emissivity, from 30 cm distance.

2.2 Preprocessing of Pseudocolored Thermograms

Since the main purpose of this research is revealing the temperature differences after the thermoregulation experiment, we selected the common Rainbow and very uncommon Lava for pseudocoloration, instead of grayscale. The background emits radiation as well, therefore existence of a significant contrast between the hands and the background is crucial. The contrast provided by Rainbow enables the segmentation through the highest difference between the hands and the background among the other built-in pseudocolorization algorithms, yet the representation of low-radiation is questionable.

On the contrary, Lava thermograms are significantly better to identify the cold regions, nonetheless the contrast between the cold spots and the background is limited. While applying any algorithm, erroneous segmentation arises due to low contrast. Therefore the optimal scenario for segmentation and low-radiation identification is using these two thermograms, presented in Fig. 1, together.

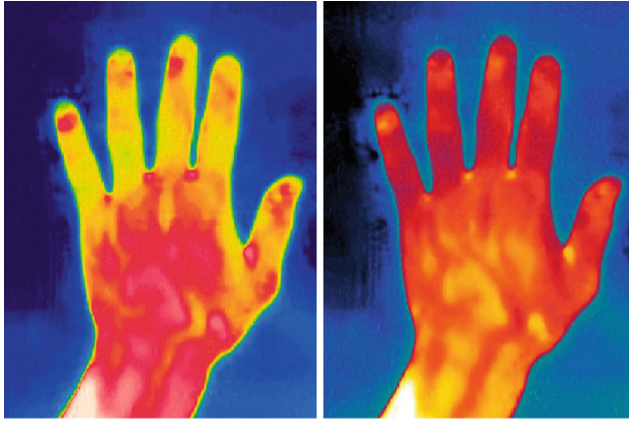


Fig. 1. Built-in pseudocolored thermograms (rainbow on the left, lava on the right) (Color figure online)

Given these facts, any pixel $p_{i,j,k}$ on the Rainbow image $R_{i,j,k}$ and $q_{i,j,k}$ on the Lava image $L_{i,j,k}$ could be represented as

$$p_{i,j,k} \in R_{i,j,k}(i = [1 : w], j = [1 : h], k = [1 : 3]) \quad (1)$$

$$q_{i,j,k} \in L_{i,j,k}(i = [1 : w], j = [1 : h], k = [1 : 3]) \quad (2)$$

where w is the width, h is the height of the image; i is the row number, j is the column number, k is the RGB color channel; 1 for Red, 2 for Green and 3 for Blue.

2.3 Hand Segmentation

Through the high contrast that the Rainbow image provides, it is easier to segment the hands by hard k-means, clustering the pixels into two sets, despite numerous alternatives. When the clusters are predefined, each pixel on the color channels would be assigned to existing clusters which is a process of fuzzy k-means and yet still classification is still crisp. Therefore, the objective function for $c = 2$ clusters could be defined as:

$$J = \sum_{a=1}^c \sum_{b=1}^{i,j,k} u_{a,b}^m d_{ab} \quad (3)$$

for the cluster C_b with memberships $u_{a,b}$ where i,j,k represents the data points and d_{ab} is the squared Euclidean distance between data points X_b and clusters C_b so that $u_{a,b}$ should satisfy:

$$\sum_{b=1}^c u_{a,b} = 1 \tag{4}$$

Therefore $u_{a,b}$ is updated by each iteration while $d_{ab} > \eta$ as follows:

$$u_{a,b} = \left((d_{ab})^{1/(m-1)} \sum_{s=1}^c \left(\frac{1}{d_{as}} \right)^{1/(m-1)} \right)^{-1} \tag{5}$$

When $d_{ab} \leq \eta$, $u_{a,b} = 1$ and $u_{a,s} = 0$ where $\eta = 10^{-6}$ for this project. Finally the cluster representatives, which correspond the palette values for red green and blue, could be calculated as follows:

$$[C_b] = \sum_{a=1}^N u_{ab}^m X_b / \sum_{a=1}^N u_{ab}^m \tag{6}$$

For the Rainbow image presented in Fig. 1 left, the cluster representatives as the palette matrix is computed in uint8, is a number format of 8-bit integer from 0 to 255 as follows

$$[C_b] = \begin{bmatrix} 1 & 54 & 140 \\ 246 & 121 & 44 \end{bmatrix} \tag{7}$$

however this matrix is not applicable in its present form that produced $\bar{R}_{i,j,k}$ as seen in Fig. 2 left. Therefore the image is reproduced as binary image $B_{i,j}$ by



Fig. 2. Segmented hands from $R_{i,j,k}$ ($\bar{R}_{i,j,k}$ generated by calculated color palette on the left, $B_{i,j}$ by manipulated binary color palette on the right) (Color figure online)

$$[B_b] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \tag{8}$$

The binary image, B_{ij} is mandatory when dealing with two kinds of thermograms as a decisive matrix. The final segmented Lava image is found by dot-product $\bar{L}_{i,j,k} = L_{i,j,k} \cdot B_{ij}$ as in Fig. 3.

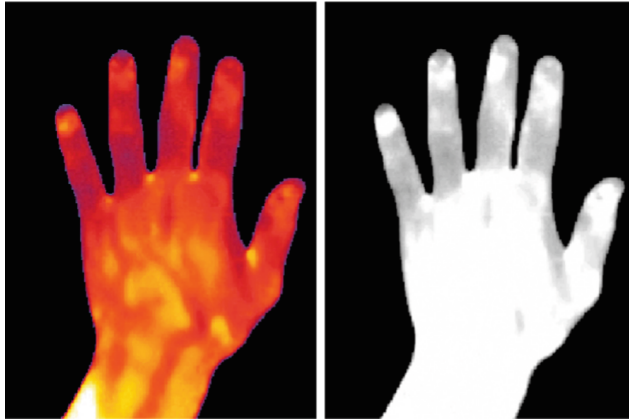


Fig. 3. Segmented hands $\bar{L}_{i,j,k}$ on the left with corresponding red color channel $\bar{L}_{i,j,1}$, on the right (Color figure online)

In lava image, as we already stated that the red channel represents the highest radiation despite all other pseudocolored thermograms. For instance in rainbow colormap the higher radiation is represented with white pixels formed by maximum values of red green and blue channels. Although this feature could be found in grayscale thermograms, the veins would have been extracted as well which will jeopardize the main purpose.

3 The Main Experiment

The red channel of the segmented lava image in Fig. 3 right, includes low-radiated regions after the thermoregulation experiment we conducted. Therefore we detected these regions by two-class fuzzy c-means to identify homogeneity after warming process. The main objective in this process is calculating threshold levels for assigning pixels to corresponding clusters.

Given the binary characteristics of $\bar{L}_{i,j,1}$, we tried to partition i,j pixels into 3 clusters by minimizing:

$$Q = \sum_{a=1}^c \sum_{b=1}^{ij} u_{a,b}^m \|X_b - v_a\|^2 \tag{9}$$

where $u_{a,b}$ represents the degree of membership of pixel a to cluster b , $m = 2$ is weighting factor, X_b is the data value, v_a is the cluster center to satisfy the Eqs. (4) and (5) again. Considering the number of clusters, the thresholds are defines as:

$$T_1 = \frac{\max(C_1) + \min(C_2)}{2} \tag{10}$$

$$T_2 = \frac{\max(C_2) + \min(C_3)}{2} \tag{11}$$

The pixels $f_{ij}^{T_1}$ and $f_{ij}^{T_2}$ in the final binary images $F_{ij}^{T_1}$ and $F_{ij}^{T_2}$ are reconstructed by

$$f_{ij}^{T_1} \in F_{ij}^{T_1} = \begin{cases} 1 & | f_{ij}^{T_1} \geq T_1 \\ 0 & | O/W \end{cases} \tag{12}$$

$$f_{ij}^{T_2} \in F_{ij}^{T_2} = \begin{cases} 1 & | f_{ij}^{T_2} \geq T_2 \\ 0 & | O/W \end{cases} \tag{13}$$

and the following images are acquired (Fig. 4).

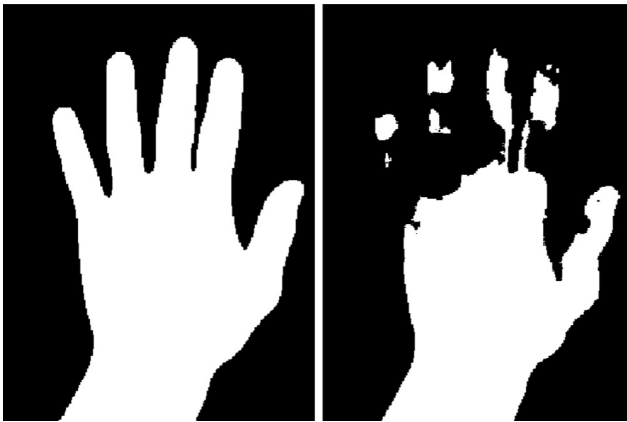



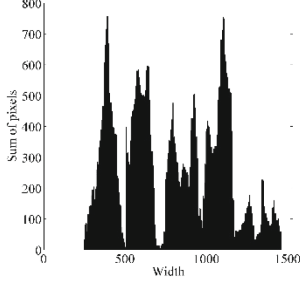

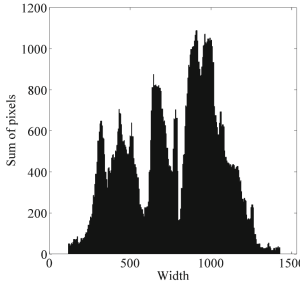

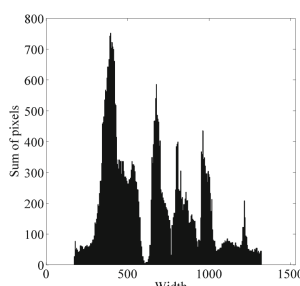
Fig. 4. Clustered low-radiated regions according to $T_1 = 0.37$ and $T_2 = 0.86$. $f_{ij}^{T_1}$ on the left, $f_{ij}^{T_2}$ on the right.

The final image of this experiment is however achieved by: $f_{i,j} \in F_{i,j} = f_{i,j}^{T_1} - f_{i,j}^{T_2}$ to make the low-radiated regions white.

4 Results

After the final images are acquired, the low-radiation regions are colored in white, therefore the vertical summation of the matrix provides the necessary information about temperature distribution after thermoregulation experiment (Table 2).

Table 2. Experimental results

Final Images $F_{i,j}$	Vertical Summation
	
	
	

It is seen that the summation diagram in first line including the experiment mentioned throughout the paper is totally balanced; while in the second line, the distribution is skewed to the right; in the third, to the left. Therefore the purpose of this paper

is accomplished by differentiating the vasospasm or vasoconstriction situations or more generally thermoregulation irregularities from the normal and healthy thermoregulation results.

5 Conclusions and Discussions

Vasospasm in hands, acute or chronic, is an important disorder causing extreme vasoconstriction that reduces blood flow to the fingers. Due to vascular changes, it could also cause cyanosis and discoloration in severe cases. Reduction in blood flow could be diagnosed by thermal imaging by analyzing the thermograms after the thermoregulation experiment. Cold water immersion narrows down the hand vessels and subsequently the vasodilatation process in healthy subjects creates a homogeneous pattern of high and low temperature regions.

We therefore presented a methodology using fuzzy k-means for segmentation; fuzzy c-means for identification of the distribution of low-radiation. We firstly made the thermoregulation and recorded one image per experiment by two types of pseudocolorization: Rainbow and Lava. High contrast in Rainbow images enabled very smooth segmentation by hard k-means method. The binary image as the output of this process is used as a multiplier to achieve the segmented Lava image. Subsequently, fuzzy c-means clustering is applied to the Lava images to identify cold regions and by summation the distribution of the cold region is achieved.

The preliminary results of utilization of two kinds of fuzzy algorithms seem promising according to the summation diagrams though we haven't generated any analytical outcome for automatic identification. Furthermore, we methodology we put forward is new for medicine and the scientists dealing with thermogram analysis.

On the other hand, the algorithms could be diversified by applying to different types of pseudocolored thermograms for identifying high radiation instead. In some research papers [22], it is stated that vasospasms could cause excessive hyperemia after cold water immersion which enlarges blood vessels instead of vasoconstriction. Moreover, if it is necessary to segment the hands in visible spectrum, the adaptive method we proposed in [23] could be useful.

Acknowledgment. The work and the contribution were supported by the SPEV project "Smart Solutions in Ubiquitous Computing Environments", University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic. We are also grateful for the support of Ph.D. students of our team (Ayca Kirimtat and Pavel Blazek) in consultations regarding application aspects.

References

1. Das, S., Maiti, A.: Acrocyanosis: an overview. *Indian J. Dermatol.* **58**(6), 417–420 (2013)
2. Parker, L.K., Ponte, C., Howell, K.J., Ong, V.H., Denton, C.P., Schreiber, B.E.: Clinical features and management of erythromelalgia: long term follow-up of 46 cases. *Clin. Exp. Rheumatol.* **35**(1), 80–84 (2017)

3. Boskovski, M.T., Thomson, J.G.: Acroparesthesia and carpal tunnel syndrome: a historical perspective. *J. Hand Surg.* **39**(9), 1813–1821 (2014)
4. Carter, P.M., Hollinshead, P.A., Desmond, J.S.: Hypothenar hammer syndrome: case report and review. *J. Emerg. Med.* **45**(1), 22–25 (2013)
5. Wigley, F.M.: Raynaud's phenomenon. *New Engl. J. Med.* **347**(13), 1001–1008 (2002)
6. Mosier, B.A., Hughes, T.B.: Recurrent carpal tunnel syndrome. *Hand Clin.* **29**(3), 427–434 (2013)
7. Lim, M.J., Kwon, S.R., Jung, K.H., Joo, K., Park, S.G., Park, W.: Digital thermography of the fingers and toes in Raynaud's phenomenon. *J. Korean Med. Sci.* **29**(4), 502–506 (2014)
8. Kolesov, S.N.: Thermal-vision diagnosis of Raynaud's syndrome and its stages. *J. Opt. Technol.* **82**(7), 478–486 (2015)
9. Ammer, K.: The sensitivity of infrared imaging for diagnosing Raynaud's phenomenon is dependent on the method of temperature extraction from thermal images. In: *Infrared Imaging*, pp. 161–166. IOP Publishing (2015)
10. Herrick, A.L.: The pathogenesis, diagnosis and treatment of Raynaud phenomenon. *Nat. Rev. Rheumatol.* **8**(8), 469–479 (2012)
11. Pauling, J.D., Shipley, J.A., Harris, N.D., McHugh, N.J.: Use of infrared thermography as an endpoint in therapeutic trials of Raynaud's phenomenon and systemic sclerosis. *Clin. Exp. Rheumatol.-Incl Suppl.* **30**(2), 103–115 (2012)
12. Roustit, M., Blaise, S., Millet, C., Cracowski, J.L.: Impaired transient vasodilation and increased vasoconstriction to digital local cooling in primary Raynaud's phenomenon. *Am. J. Physiol.-Heart Circulat. Physiol.* **301**(2), 324–330 (2011)
13. Ring, F., Jung, A., Žuber, J.: *Infrared Imaging: A Casebook in Clinical Medicine*. IOP Publishing, Bristol (2015)
14. Alpar, O., Krejcar, O.: Detection of Raynaud's phenomenon by thermographic testing for finger thermoregulation. In: *Intelligent Information and Database Systems*, pp. 475–484 (2017)
15. Snehalatha, U., Anburajan, M., Teena, T., Venkatraman, B., Menaka, M., Raj, B.: Thermal image analysis and segmentation of hand in evaluation of rheumatoid arthritis. In: *2012 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–6. IEEE (2012)
16. Mohiyuddin, N., Dhage, P., Warhade, K.K.: Segmentation of thermal images for evaluation of rheumatoid arthritis disease. *Int. J. Emerg. Eng. Res. Technol.* **2**(4), 54–63 (2014)
17. Antony, S.J.S., Ravi, S.: Breast cancer detection on thermogram at preliminary stage by using fuzzy inferences system. *J. Theor. Appl. Inf. Technol.* **68**(3), 705–715 (2014)
18. Ring, E.F.J., Ammer, K.: Infrared thermal imaging in medicine. *Physiol. Meas.* **33**(3), 33–46 (2012)
19. Dinsdale, G., Herrick, A.L.: Vascular diagnostics for Raynaud's phenomenon. *J. Vasc. Diagn.* **2**, 127–139 (2014)
20. Ismail, E., Orlando, G., Corradini, M.L., Amerio, P., Romani, G.L., Merla, A.: Differential diagnosis of Raynaud's phenomenon based on modeling of finger thermoregulation. *Physiol. Meas.* **35**(4), 703–716 (2014)
21. Foerster, J., Wittstock, S., Fleischanderl, S., Storch, A., Riemekasten, G., Hochmuth, O., Worm, M.: Infrared-monitored cold response in the assessment of Raynaud's phenomenon. *Clin. Exp. Dermatol.* **31**(1), 6–12 (2006)
22. Daanen, H.: Cold-induced vasodilation. *Eur. J. Appl. Physiol.* **105**(4), 663–664 (2009)
23. Alpar, O., Krejcar, O.: Dorsal hand recognition through adaptive YCbCr Imaging technique. In: Nguyen, N.-T., Manolopoulos, Y., Iliadis, L., Trawiński, B. (eds.) *ICCCI 2016 Part II. LNCS (LNAI)*, vol. 9876, pp. 262–270. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-45246-3_25



Medical Image Classification with Hand-Designed or Machine-Designed Texture Descriptors: A Performance Evaluation

Joke A. Badejo¹(✉), Emmanuel Adetiba^{1,2,3},
Adekunle Akinrinmade¹, and Matthew B. Akanle^{1,2}

¹ Department of Electrical and Information Engineering, Covenant University,
Canaanland, Ota, Nigeria

{joke.badejo, emmanuel.adetiba,
bola.akanle}@covenantuniversity.edu.ng,
adekunleakinrinmade@gmail.com

² Center for Systems and Information Services (CSIS), Covenant University,
Canaanland, Ota, Nigeria

³ HRA, Institute for Systems Science, Durban University of Technology,
P.O. Box 1334, Durban, South Africa

Abstract. Accurate diagnosis and early detection of various disease conditions are key to improving living conditions in any community. The existing framework for medical image classification depends largely on advanced digital image processing and machine (deep) learning techniques for significant improvement. In this paper, the performance of traditional hand-designed texture descriptors within the image-based learning paradigm is evaluated in comparison with machine-designed descriptors (extracted from pre-trained Convolution Neural Networks). Performance is evaluated, with respect to speed, accuracy and storage requirements, based on four popular medical image datasets. The experiments reveal an increased accuracy with machine-designed descriptors in most cases, though at a higher computational cost. It is therefore necessary to consider other parameters for tradeoff depending on the application being considered.

Keywords: Medical image classification · Deep learning
Convolution Neural Network · Texture descriptors

1 Introduction

The development of computer-aided diagnosis, CADx systems for early detection and accurate diagnosis of deadly diseases has gained a lot of attention in the research community. CADx systems (also called computer-aided detection, CAde) are systems that assist doctors in the interpretation of medical images. Imaging techniques in X-ray, MRI, and ultrasound diagnostics yield a great deal of information that the radiologist, pathologist or other medical professional have to analyze and evaluate comprehensively in a short time. CAD systems process digital images for typical appearances and

to highlight conspicuous sections, such as possible diseases, in order to offer input to support a decision taken by the professional [1–5].

These systems have facilitated significant improvements in the medical sector. For example, CAD systems have been employed to assist physicians in the early detection of breast cancers on mammograms by improved diagnoses of malignant and benign clustered micro-calcifications for over three decades. Traditionally, only 15–30% of breast biopsies for non-palpable lesions result in the discovery of breast malignancies, while the majority of these procedures are performed on benign lesions [6, 7]. CAD perform at an equivalent or higher accuracy compared to radiologists [8] and helps to resolve benign-outcome biopsy (i.e., false-positive mammograms).

CAD systems decrease observational oversights and thus the false negative rates of physician’s interpretations and serves as a “second opinion” in assisting radiologists’ image interpretations. Since the number of imaging studies grows steadily, and the number of images per study grows drastically, the use of a computer rather than a second human observer has the advantage of not increasing the demands on the radiologist (or trained observer) pool [9, 10].

Modern CAD systems are powered by the trio combination of digital image processing, machine (or deep) learning techniques and computing technology. Models are usually based on traditional hand-designed texture descriptors such as Local Binary Patterns (LBP), Local Phase Quantization (LPQ) and their variants [1, 4–6]; and lately on descriptors from convolutional neural networks, CNN and their variants [11–14].

Spanhol et al. performed classification of BC histopathological images of the BreakHis dataset by adapting AlexNet, an existing non-handcrafted architecture that was used for classifying color images of objects by extracting patches to train the CNN and for final classification. Their work shows that the results obtained using CNN were more accurate compared to traditional machine learning models trained on the same dataset but with state of the art texture descriptors [15].

Gao et al. presented a paper systematically comparing classifications utilizing deep convolutional neural networks (CNNs) with the well-established image classification models in the literature, namely bag-of-features (BoF) and Fisher Vector (FV) on HEP-2 cell images. Their results were also divided with CNN-based framework performing slightly better with augmentation and handcrafted framework performing better otherwise [16].

Convolutional Neural Networks represented a major breakthrough in computer vision, having significantly improved the state-of-the-art in many applications. Originally developed for object and scene classification, the approach proved effective in other domains as well, for example face recognition [17]. It is, however, still a subject of debate whether this paradigm is amenable to being successfully applied to fine-grained images. The deep learning strategies have not yet been sufficiently explored, thus we might expect new relevant results using such approaches.

The works of Liu et al. [18] considered other important factors in their comparisons. Their experiments were designed to measure robustness against different classification challenges, including changes in rotation, scale, illumination, viewpoint, number of classes, different types of image degradation, and computational complexity. Their findings showed that the best overall performance was obtained for handcrafted texture descriptors. However, in cases where the textures had large variations in appearance, non-handcrafted methods performed best but at the cost of very high computational complexity.

More insight was obtained from the works done by Bello-Cerezo et al. They carried out experiments on ten datasets using eleven hand-designed local image descriptors and nine sets of CNN-based features on grey-scale images. Their findings show that the results were divided. CNN-based features seemed to perform better when the intra-class variability was high while the LBP variants yielded better results for lower intra-class variability [17]. A similar result to the above was also observed in the works of Hertel et al. [19]. They obtained marginal improvement using non-handcrafted architecture on datasets; MNIST, CIFAR-10, CIFAR-100. The typical long time needed to train such deep systems was reduced by reusing a previously trained network on the ILSVRC-12 dataset.

For these reasons, it is necessary to compare, to facilitate better understanding of the state of the art LBP-based approaches and provide guidelines for practitioners to decide which features to use according to their requirements [18]. CNNs also have the advantage of being able to learn generic feature extractors that can be used for different tasks [19] and have demonstrated excellent adaptability across different datasets, which is highly desirable for classification under varying laboratory settings [16].

This paper evaluates the performance of hand-designed and machine-designed descriptors, considering speed, accuracy, memory requirements and rate of early diagnosis where applicable. The evaluation is carried out on four popular medical image datasets: 2D-HeLa dataset for cell phenotype image classification [20], Pap smear datasets for detecting abnormal smear cells [21], Chinese Hamster Ovary Cells dataset [22] and Breast Cancer Histopathological Images, BreaKHis [15].

2 Materials and Methods

2.1 Medical Image Datasets

The images used in this paper are selected from four medical image datasets as briefly described in this section.

Pap Smear Dataset: contains 917 images, obtained by a digital camera and a microscope, of cells scraped from the cervix to diagnose cancer. The cells provided are categorized into two broad classes and seven subclasses: 242 normal cells (intermediate squamous, superficial and columnar epithelial) and 675 abnormal cells (mild dysplasias, moderate dysplasia, severe dysplasia and squamous cell carcinoma in situ intermediate) [21].

Chinese Hamster Ovary Cells Dataset: contains 327 fluorescent microscopic images of the ovary cells of the Chinese Hamster. The cell images are grouped into five classes (giantin, Hoechst, lamp2, nop4 and tubulin) [20].

2D HeLa Dataset: contains 862 fluorescent microscopy single-cells images in ten staining classes: Actin, Nucleus, Endosome, ER, Golgi Giantin, Golgi GPP130, Lyso-some, Microtubules, Mitochondria, and Nucleolus. Images are useful for classifying protein subcellular patterns in order to understand the functions and behavior of the protein [20].

Breast Cancer Histopathological Images Dataset: contains 7909 microscopic biopsy images of breast tumors in four magnification factors: 40x (1995), 100x (2081), 200x (2013) and 400x (1820). The images are grouped into two classes and eight subclasses: benign tumors (adenosis, fibroadenoma, phyllodes tumor and tubular adenoma) and malignant tumors (ductal carcinoma, lobular carcinoma, mucinous carcinoma and papillary carcinoma) [15].

2.2 Image Texture Descriptors

Extensive survey of texture descriptors for various computer vision applications and especially medical image analysis for CAD shows the significant successes of local texture descriptors designed with signal processing techniques (hand-designed) and more recently, those learned from the images by convolutional neural networks (machine-designed) [1, 18, 23, 24]. In this paper, we evaluated how well suited the filter banks from these two broad classes of texture descriptors are, in context of these applications.

We provide a brief description of these features here.

Table 1. Average training time per dataset (*seconds*), for each classifier

Texture descriptor	#Features	Classifier	Datasets						
			PS	CHO	HeLa	BH40x	BH100x	BH200x	BH400x
LBP ^{riu2}	28	SVM	5.04	2.13	5.66	1.64	2.14	1.29	1.64
		NNC	1.13	0.93	1.02	1.12	1.13	2.37	1.17
LTP ^{riu2}	20	SVM	3.66	1.62	5.65	1.65	2.26	2.12	2.12
		NNC	0.83	0.91	1.06	1.26	1.89	1.16	1.12
CLBP_M ^{riu2}	28	SVM	3.64	2.13	5.66	2.14	2.13	2.16	2.12
		NNC	0.84	0.92	1.12	1.12	1.14	1.15	1.12
CLBP_S ^{riu2}	28	SVM	3.65	2.12	6.22	1.12	2.13	2.14	1.64
		NNC	0.88	0.93	1.07	4.67	1.15	1.14	1.14
LPQ	256	SVM	6.92	2.64	9.72	4.66	5.68	4.80	4.72
		NNC	2.2	1.14	2.68	4.62	5.18	4.67	4.15
RICLBP	272	SVM	6.7	2.65	9.22	4.68	5.68	5.25	4.70
		NNC	2.17	1.13	2.17	4.72	5.18	4.69	4.18
AlexNet ('fc7')	4096	SVM	90.27	29.42	104.5	96.76	104.36	93.11	85.58
		NNC	30.78	19.33	30.47	84.43	78.70	74.14	65.49
<i>#Training Images per Dataset</i>			<i>642</i>	<i>229</i>	<i>603</i>	<i>1397</i>	<i>1457</i>	<i>1409</i>	<i>1274</i>
<i>#Class per Dataset</i>			<i>7</i>	<i>5</i>	<i>10</i>	<i>2</i>	<i>2</i>	<i>2</i>	<i>2</i>

2.2.1 Local Binary Pattern, LBP

LBP [25] is a texture descriptor that encodes the signed magnitude difference between a center pixel and its adjacent or neighboring pixels in two levels. Given an image intensity value I of a pixel p , at position (x, y) ,

$$LBP_{r,N}(p) = \sum_{i=0}^{N-1} \text{sgn}(I(p + \Delta r_i) - I(p))2^i \quad (1)$$

$$\text{where } \text{sgn}(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (1.1)$$

N is the number of the neighbor pixels, defines LBP resolution, r is the radius of the LBP, Δr_i is defined as a unit displacement vector corresponding to spatial coordinates ($N = 8$):

$$\Delta r = \{(r, 0), (r, r), (0, r), (-r, r), (-r, 0), (-r, -r), (0, -r), (r, -r)\} \quad (1.2)$$

LBP captures both the structural and statistical properties of intrinsic texture elements and is noted for its robustness to illumination changes in an image. In this paper, we consider the subset of LBPs termed rotation invariant uniform patterns, *riu2*.

2.2.2 Local Ternary Pattern, LTP

LTP [27], also a variant of LBP where the number of quantization levels was increased from 2 to 3. It encodes the gray level difference around zero to address the problem of noise sensitivity in near-uniform image regions.

2.2.3 Completed Local Binary Pattern, CLBP

CLBP [29], is a variant of LBP where the differences between a center pixel and its neighbors are decomposed into the sign, CLBP_S, the magnitude components, CLBP_M and the binary-encoded center pixel gray values, CLBP_C.

2.2.4 Local Phase Quantization, LPQ

LPQ [28] as inspired by LBP, is a blur-tolerant texture descriptor. It quantizes the 2-D short-term Fourier transform phase in the local neighborhood, as against the pixel gray values in LBPs, to obtain a histogram of the encoded labels.

2.2.5 Rotation Invariant Co-occurrence of adjacent LBPs, RICLBP

RICLBP [30] captures the structural relationships between two adjacent uniform LBPs in a local neighborhood at different unit normal direction to the center pixel. The spatial distance between the LBPs can be varied to have a multi-scale texture representation.

2.2.6 Activations from pre-trained CNN

CNN [26] as proposed in previous works like [17], [19], the activations (or filter banks) learnt from a pre-trained CNN can be used to adopted as generic texture descriptor. In this work, we used the last fully connected layer ‘fc7’ of AlexNet [26] to extract features from the images. It is important to note that the ‘machine-designed’ filter banks obtained at this layer are considered to obtain low-level features from images, different from image-specific features.

3 Experimental Setup and Results

All the datasets used for evaluation in this work were split into 2 disjoint sets for training and testing using the 10-fold cross validation protocol, with 70% for training and 30% for testing. In our experimental setup, we adopted $(r, n) = (1, 8), (2, 16)$ for LBP, LTP, CLBP and $(r, d) = (1, 2), (2, 4)$ for RICLBP for multi-scale and multi-resolution texture modeling. The resulting histograms from each parameter pair of the operators are concatenated to obtain the feature vector. The classifiers used in this work are Nearest Neighbor Classifier, NNC ($k = 1$) with the L_2 measure and Support Machine Vector, SVM with the radial basis function, RBF. Kernel parameter is empirically set to 4.5 (LTP), 5.3 (LBP, CLBP), 16 (RICLBP, LPQ) and 64 (AlexNet).

Table 2. Correct classification accuracy, CCR

Texture descriptor	Classifier	Datasets						
		PS	CHO	HeLa	BH40x	BH100x	BH200x	BH400x
LBP ^{riu2}	SVM	43.3	76.1	75.6	79.7	80.7	75.3	76.0
	NNC	38.9	81.0	69.5	82.2	79.2	75.5	73.6
LTP ^{riu2}	SVM	48.3	78.6	72.7	75.6	75.4	73.6	72.5
	NNC	40.8	87.8	65.9	75.7	74.9	73.4	68.1
CLBP_M ^{riu2}	SVM	43.2	79.2	78.1	79.4	80.2	75.7	76.4
	NNC	31.6	80.7	72.4	81.9	78.6	76.7	73.4
CLBP_S ^{riu2}	SVM	43.3	84.7	73.8	79.4	79.3	72.5	73.2
	NNC	40.0	89.3	64.4	77.5	75.3	71.6	67.6
LPQ	SVM	47.0	78.6	70.8	91.1	90.7	86.2	84.3
	NNC	34.5	78.0	68.3	90.3	89.6	87.2	81.4
RICLBP	SVM	52.6	73.1	68.0	89.3	84.8	79.8	78.5
	NNC	41.0	84.7	63.3	87.9	82.6	80.9	76.0
AlexNet ('fc7')	SVM	61.5	77.7	63.6	86.8	85.5	86.8	87.0
	NNC	53.1	64.5	46.1	86.9	83.6	83.9	82.9

Performance evaluation of the system was measured with Average Training time (in seconds), Correct Classification Rate (CCR), Area Under Curve (AUC), and prediction speed (#observations per second) are presented in Tables 1, 2, 3, and 4 respectively. All experiments were carried out with Intel Xeon 3.1 GHz with 8 GB RAM and Matlab2017, pre-trained models of AlexNet were obtained from Mat-ConvNet [27].

A summary of the texture descriptors used in the experiments, the number of features, the average training time in seconds and the number of images in the training set and number of classes is presented in Table 1. It is noted that the training time for each dataset is directly proportional to the number of classes involved.

The results obtained from the experiments, as presented in Tables 2 and 3, varies largely with the dataset and the type of laboratory or imaging techniques used to produce the images. The Pap Smear images are low-resolution colored images

Table 3. Area Under the Curve, AUC

Texture descriptor	Classifier	Datasets						
		PS	CHO	HeLa	BH40x	BH100x	BH200x	BH400x
LBP ^{riu2}	SVM	0.87	0.99	0.99	0.88	0.86	0.82	0.79
	NNC	0.94	0.96	0.99	0.80	0.76	0.71	0.71
LTP ^{riu2}	SVM	0.96	0.99	0.99	0.80	0.81	0.77	0.72
	NNC	0.80	0.96	0.99	0.72	0.72	0.70	0.64
CLBP_M ^{riu2}	SVM	0.93	0.99	0.99	0.88	0.86	0.83	0.79
	NNC	0.70	0.96	0.99	0.79	0.75	0.73	0.70
CLBP_S ^{riu2}	SVM	0.94	0.99	0.99	0.85	0.83	0.76	0.73
	NNC	0.79	0.99	0.99	0.75	0.73	0.67	0.64
LPQ	SVM	0.97	0.94	0.99	0.96	0.96	0.93	0.90
	NNC	0.86	0.79	0.98	0.90	0.88	0.85	0.80
RICLBP	SVM	0.97	0.98	0.99	0.95	0.92	0.88	0.86
	NNC	0.77	0.96	0.99	0.87	0.80	0.78	0.73
AlexNet ('fc7')	SVM	0.99	0.95	0.99	0.92	0.91	0.94	0.93
	NNC	0.87	0.83	0.97	0.83	0.81	0.81	0.82

compared to BreakHis images, with machine-learned descriptors producing better results. With a good resolution, the handcrafted descriptors seem able to extract important primitive information from the images to differentiate a cancer tumor from non-cancerous cells even at a magnification factor of 40. On the other side, HeLa and CHO images are binary images which appears to work better with handcrafted images. Likewise, the descriptor-classifier combination was considered in our experiments. SVM appears to work significantly well with most images even for multiple categories or classes. However, we think this combination can afford the opportunity of tuning the parameters of a generic CAD system for specific application scenarios.

4 Discussion and Conclusion

The empirical evaluation in this work facilitated a better understanding of the state-of-the-art performance of LBP-based approaches and pre-trained CNNs on medical image datasets. The experiments reveal an increased accuracy with CNN-based descriptors in difficult cases, such as Pap Smear and some BreakHis images, but almost at-par results with LBP or its close variants for other images. However, considering the computational complexity of CNNs in terms of training time, prediction speed (or number of samples predicted per second), memory resources as shown in Table 1 and 4, it demands a higher cost as a tradeoff to accuracy.

With the lower computational complexity and high success rate of human-designed or handcrafted descriptors such as LBP and HoG [18, 28, 29], CAD systems designed with them appears quite promising and especially for ascertaining the present of cancer cells or not. In the real sense, CNNs are designed to solve more complex problems that involves various noise illumination, pose, image resolution/quality, blur, occlusion etc.;

Table 4. Prediction speed (\sim # of samples/second)

Texture Descriptor	Classifier	Datasets						
		PS	CHO	HeLa	BH40x	BH100x	BH200x	BH400x
LBP ^{riu2}	SVM	2400	2000	1900	9900	19000	9800	17000
	NNC	9100	3600	9200	19000	9500	8500	10000
LTP ^{riu2}	SVM	3000	2100	1900	9300	20000	20000	19000
	NNC	9100	4600	9700	9500	9700	11000	12000
CLBP_M ^{riu2}	SVM	2900	2100	1800	19000	19000	17000	18000
	NNC	9300	3900	8400	10000	9400	10000	10000
CLBP_S ^{riu2}	SVM	3100	2100	1800	19000	18000	16000	17000
	NNC	9200	3800	8300	10000	10000	9100	10000
LPQ	SVM	1300	990	910	5000	5100	5000	4500
	NNC	2100	1300	1900	1700	1600	1700	1800
RICLBP	SVM	1200	980	920	4900	4700	4500	4600
	NNC	2000	1300	2000	1600	1600	1600	1700
AlexNet ('fc7')	SVM	71	57	57	190	180	190	180
	NNC	110	61	100	93	92	94	97

and differentiate between hundreds of classes especially with higher intra-class (than inter-class) variability [14, 19, 30, 31]. Thus, in cases where it is important to determine the exact type of cancer in a sample, CNNs might be most appropriate. This provides guidelines for practitioners to decide which features to use according to their requirements - speed, laboratory/imaging techniques, training time, computational resources/cost; in developing suitable systems for their application scenarios.

The need to improve on the existing process model for CAD systems in our society cannot be over-emphasized, as the target is to prevent as many unpleasant occurrences as possible and within the lowest budget. A quick consideration will be the feasibility of having these CAD systems as handheld or embedded devices in the new paradigm of 'smart and intelligent' systems. Next, it is also important to consider the reduction in the overall cost of the CAD system for deployment in low resource settings where the medical professionals and experts are almost not available.

The CNN considered in this work is limited to AlexNet; we are still exploring the very-deep network, VGG and other variants to find suitable models even for larger and various datasets, such as Hep-2, for the task at hand. The process of training a deep CNN can be complicated by overfitting and convergence issues, requiring repetitive adjustments in the architecture or learning parameters of the network to ensure that all layers are learning with comparable speed. Machine-learned filter banks, obtained at the convolution layers of CNN, also give insights into essential image details required for better performance. It is important to note that the 'machine-designed' filter banks obtained at these layers are considered to obtain low-level features from images. Future works will be to evaluate other concepts such as the Bag of Visual Words, BoVW pipeline for successful hand-crafted descriptors; develop a model to combine LBP variants with deep convolutional networks or even as building blocks of novel cascaded deep convolutional network architectures as reviewed in [18, 24].

Acknowledgements. This research was sponsored under the Centre for Research Innovation and Development Research Grant of Covenant University. The authors who shared their MATLAB code and toolboxes for LBP, LTP, LPQ, CLBP, RICLBP and MatConvNet are appreciated.


References

1. Nanni, L., Lumini, A., Brahnam, S.: Local binary patterns variants as texture descriptors for medical image analysis. *Artif. Intell. Med.* **49**(2), 117–125 (2010)
2. Lumini, A., Nanni, L., Brahnam, S.: Multilayer descriptors for medical image classification. *Comput. Biol. Med.* **72**, 239–247 (2016)
3. Verma, B., McLeod, P., Klevansky, A.: Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer. *Expert Syst. Appl.* **37**(4), 3344–3351 (2010)
4. Moura, D.C., López, M.A.G.: An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. *Int. J. Comput. Assist. Radiol. Surg.* **8**(4), 561–574 (2013)
5. Liu, D., Wang, S., Huang, D., Deng, G., Zeng, F., Chen, H.: Medical image classification using spatial adjacent histogram based on adaptive local binary patterns. *Comput. Biol. Med.* **72**, 185–200 (2016)
6. Kopans, D.B.: The positive predictive value of mammography. *Am. J. Roentgenol.* **158**, 521–526 (1992)
7. Knutzen, A.M., Gisvold, J.J.: Likelihood of malignant disease for various categories of mammographically detected, nonpalpable breast lesions. In: *Mayo Clinic Proceedings* (1993)
8. Jiang, Y., et al.: Malignant and benign clustered microcalcifications: automated feature analysis and classification. *Radiology* **198**, 671–678 (1996)
9. Hobson, P., Lovell, B.C., Percannella, G., Saggese, A., Vento, M., Wiliem, A.: HEp-2 staining pattern recognition at cell and specimen levels: datasets, algorithms and results. *Pattern Recognit. Lett.* **82**, 12–22 (2016)
10. Rubin, G.D.: Data explosion: the challenge of multidetector-row CT. *Eur. J. Radiol.* **2**, 74–80 (2000)
11. Shen, W., Zhou, M., Yang, F., Yang, C., Tian, J.: Multi-scale convolutional neural networks for lung nodule classification. In: Ourselin, S., Alexander, D.C., Westin, C.-F., Cardoso, M. J. (eds.) *IPMI 2015*. LNCS, vol. 9123, pp. 588–599. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19992-4_46
12. Tajbakhsh, N., et al.: Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imaging* **35**(5), 1299–1312 (2016)
13. Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D.D., Chen, M.: Medical image classification with convolutional neural network. In: 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), pp. 844–848 (2014)
14. Shin, H.C., et al.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**(5), 1285–1298 (2016)
15. Spanhol, F.A., Oliveira, L.S., Petitjean, C., Heutte, L.: A dataset for breast cancer histopathological image classification. *IEEE Trans. Biomed. Eng.* **63**, 1455–1462 (2016)
16. Gao, Z., Wang, L., Zhou, L., Zhang, J.: HEp-2 cell image classification with deep convolutional neural networks. *IEEE J. Biomed. Heal. Inform.* **21**(2), 416–428 (2017)

17. Bello-Cerezo, R., Bianconi, F., Cascianelli, S., Fravolini, M.L., di Maria, F., Smeraldi, F.: Hand-designed local image descriptors vs. off-the-shelf CNN-based features for texture classification: an experimental comparison. In: De Pietro, G., Gallo, L., Howlett, R.J., Jain, L.C. (eds.) KES-IIMSS 2017. SIST, vol. 76, pp. 1–10. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-59480-4_1
18. Liu, L., Fieguth, P., Guo, Y., Wang, X., Pietikäinen, M.: Local binary features for texture classification: taxonomy and experimental study. *Pattern Recognit.* **62**, 135–160 (2017)
19. Hertel, L., Barth, E., Kaster, T., Martinetz, T.: Deep convolutional neural networks as generic feature extractors. In: Proceedings of the International Joint Conference on Neural Networks, vol. 2015, September (2015)
20. Chebira, A., et al.: A multiresolution approach to automated classification of protein subcellular location images. *BMC Bioinform.* **8**, 210 (2007)
21. Jantzen, J., Norup, J., Dounias, G., Bjerregaard, B.: Pap-smear benchmark data for pattern classification. In: Proceedings of NiSIS 2005 Nature Inspired Smart Information System, pp. 1–9 (2005)
22. Bolland, M.V., Murphy, R.F.: A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics* **17**(12), 1213–1223 (2001)
23. Nanni, L., Ghidoni, S., Brahnam, S.: Handcrafted vs. non-handcrafted features for computer vision classification. *Pattern Recognit.* **71**, 158–172 (2017)
24. Liu, L., Chen, J., Fieguth, P., Zhao, G., Chellappa, R., Pietikäinen, M.: A survey of recent advances in texture representation. *arXiv Preprint arXiv:1801.10324* (2018)
25. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
26. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems vol. 60, no. 6, pp. 84–90 (2012)
27. Vedaldi, A., Lenc, K.: Convolutional neural networks for MATLAB (2014)
28. Adetiba, E., Olugbara, O.O.: Lung cancer prediction using neural network ensemble with histogram of oriented gradient genomic features. *Sci. World J.* **2015**, 17p (2015)
29. Adetiba, E., Olugbara, O.O.: Improved classification of lung cancer using radial basis function neural network with affine transforms of voss representation. *PLoS ONE* **10**(12), e0143542 (2015)
30. Cimpoi, M., Maji, S., Kokkinos, I., Vedaldi, A.: Deep filter banks for texture recognition, description, and segmentation. *Int. J. Comput. Vis.* **118**(1), 65–94 (2016)
31. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks, pp. 1097–1105 (2012)



Classification of Breast Cancer Histopathological Images Using KAZE Features

Daniel Sanchez-Morillo^{1,2}✉ , Jesús González³,
Marcial García-Rojo^{1,4}, and Julio Ortega³

¹ Biomedical Engineering and Telemedicine Research Group,
University of Cadiz, Puerto Real, Cadiz, Spain
daniel.morillo@uca.es,
marcial.garcia.sspa@juntadeandalucia.es

² Department of Automation,
Electronics and Computers and Network Architecture, University of Cadiz,
Puerto Real, Cadiz, Spain

³ Department of Computer Architecture and Technology,
CITIC, University of Granada, Granada, Spain
{jesusgonzalez, jortega}@ugr.es

⁴ University Hospital Puerta del Mar of Cadiz Cádiz, Cadiz, Spain

Abstract. Breast cancer (BC) is a public health problem of first importance, being the second most common cancer worldwide. BC represents 30.4% of all new cancer cases in the European female population. The diagnosis and differential diagnosis of BC is based on the clinical presentations, physical examinations combined with imaging studies, and confirmed by histopathologic findings. Pathologists' examination is a time-consuming analysis, susceptible to an interpretation bias mainly caused by the experience of the pathologist and the decrease of attention due to fatigue. Currently, computer-aided detection and diagnosis techniques applied to digital images are assisting the specialists.

In this work, the performance of a pattern recognition system based on KAZE features in combination with Bag-of-Features (BOF) to discriminate between benign and malignant tumours is evaluated on the BreakHis database (7909 images). During the training stage, KAZE keypoints are extracted for every image in the training set. Keypoints are mapped into a histogram vector using K-means clustering. This histogram represents the input to build a binary SVM classifier. In the testing stage, the KAZE keypoints are extracted for every image in the test set, and fed into the cluster model to map them into a histogram vector. This vector is finally fed into the binary SVM training classifier to classify the image.

The experimental evaluation shows the feasibility and effectiveness, in terms of classification accuracy, of the proposed scheme for the binary classification of breast cancer histopathological images with a low magnification factor.

Keywords: Breast cancer · Histopathology · Image classification
Medical imaging · KAZE · Bag of features

1 Introduction

Cancer is a public health concern and a central priority of health policy in the world. The number of cancer cases around the world is expected to increase to 24 million by 2035 [1]. Breast cancer (BC) is the second most common cancer worldwide accounting for 12% of the total of all cancers [2]. In women, BC is the most frequently diagnosed neoplasm with 30.4% of all new cancer cases in the female population in Europe [3]. BC represents the cause of death in more than 6% of the total number of cancer deaths [4]. In this regards, BC is presented as a public health problem because of its negative impact on physical, mental, social, and emotional health of women, and hence on their quality of life.

The diagnosis of BC is based on the clinical examination, combined with imaging (mammography, ultrasound and eventually magnetic resonance imaging), and confirmed by pathologist assessment [5].

The pathologist performs final diagnosis by visual inspection of histological samples taken during a breast tumour excision or biopsy. Histopathology analysis includes the tissue preparation process with formalin fixation and paraffin embedding and the staining to highlight structures of interest like nuclei and cytoplasm. The most common staining techniques currently used are immunohistochemistry (IHC) and haematoxylin and eosin (H&E). H&E has been used for decades as standard staining protocol. H&E dyes the nuclei blue/purple and other structures such as cytoplasm and stroma pink [6]. Processed tissue sections are mounted on glass slides and digitized using digital cameras and whole slide imaging (WSI) scanners. Slide scanning is commonly performed at 20x or 40x magnification factor with an approximated spatial resolution of 0.50 $\mu\text{m}/\text{pixel}$ and 0.25 $\mu\text{m}/\text{pixel}$, respectively [6].

During the histopathological image analysis, the pathologist uses several techniques and methods for obtaining quantitative data from images including size of cells, disproportionate number of cells, and tissue abnormalities [7]. It is a high time-consuming task that can be under interpretation bias inducted by different factors such as the experience of the pathologist and the decrease of attention due to fatigue [8].

Computer aided diagnosis (CAD) has become a research trend to assist the pathologist in making an accurate diagnosis and, to overcome the limitations and bias of human specialists. Different analysis techniques for BC histopathological image analysis have been proposed over the past ten years. However, the performance of most approaches has been assessed using small private datasets, making difficult a fair comparison. Large and public datasets have been lacking to the research community. In this study, the BreakHis dataset, recently introduced by Spanhol *et al.* [9] was used. The database includes 7909 breast histopathological images from 82 patients what enables the objective comparison of different strategies, approaches and computerized image analysis algorithms [6].

Common CAD approaches are based either on image segmentation to extract structure properties [10, 11], texture representation [9, 12, 13] or new deep learning methods [14, 15]. Approaches based on image segmentation include detection and segmentation of cell nuclei, features extraction and classification. Nuclei detection techniques are challenging and often suffer over-segmentation, do not perform well for overlapping cells, present poor convergence and stability issues, require prior knowledge and present a high computational complexity [10]. Deep learning methods requires a long training time and high computational resources. This work focuses on the used local features, a texture-based method.

Local features represent the texture in an image patch, and better handle scale changes and rotation. Among the applications of local features are image registration, object detection and classification, and tracking. Scale-invariant feature transform (SIFT) [16], speeded up robust features (SURF) [17], orientated BRIEF rotated FAST (ORB) [18], local binary patterns (LBP) [19], binary robust invariant scalable keypoints (BRISK) [20], histograms of oriented gradient (HOG) [21], maximally stable extremal regions (MSER) [22] and fast retina keypoint (FREAK) [23] are representative of local descriptors.

Recently, a new multiscale 2D local feature detection and description algorithm in nonlinear scale spaces called KAZE has been described [24]. KAZE detects and describes features in nonlinear scale spaces. KAZE features are used in a bag-of-features framework. KAZE features have been used very recently for feature tracking [25], image recognition [26], adaptive human tracking [27], and object detection [28], among other applications.

This work aims at evaluating the performance of a pattern recognition system for the automated classification of histopathological images of breast cancer using the BreakHis dataset. KAZE features are used in a bag-of-features approach for efficient and accurate discrimination between benign and malignant tumours. The paper is structured as follows. Section 2 details the dataset, the feature extractor, the proposed classifier and the performance measurements used in this study. Section 3 reports and describes the results. Finally, Sect. 4 summarizes the conclusions.

2 Methods

2.1 BreakHis Database

The Breast Cancer Histopathological Image Classification (BreakHis) [9] includes 7909 (2480 benign and 5429 malignant samples) microscopic images of breast tumour tissue collected by surgical (open) biopsy from 82 patients. Samples were prepared using the standard paraffin procedure and stained with H&E. Images were acquired using an Olympus BX-50 microscope coupled to a Samsung digital camera SCC-131AN. Different magnifying factors (40x, 100x, 200x, and 400x) were considered (Table 1).

Table 1. Distribution of images of benign and malignant classes in the BreakHis database.

Magnification	Benign	Malignant	Total
40x	652	1370	1995
100x	644	1437	2081
200x	623	1390	2013
400x	588	1232	1820

Image resolution is 700×460 pixels, with 3-channel RGB and 8-bit depth in each channel (TrueColor, 24-bit colour depth). Portable network graphics format (PNG) is provided.

Diagnosis of images in the dataset was performed by experienced pathologists and confirmed by complementary exams such as immunohistochemistry (IHC) analysis. The dataset includes four classes of benign breast tumours: adenosis (A), phyllodes tumour (PT), fibroadenoma (F), and tubular adenoma (TA); and four malignant tumours: ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary carcinoma (PC). Figures 1 and 2 show eight images with 400x magnification factors acquired from a single slide of breast tissues containing benign tumours (Fig. 1) and malignant tumours (breast cancer) (Fig. 2).

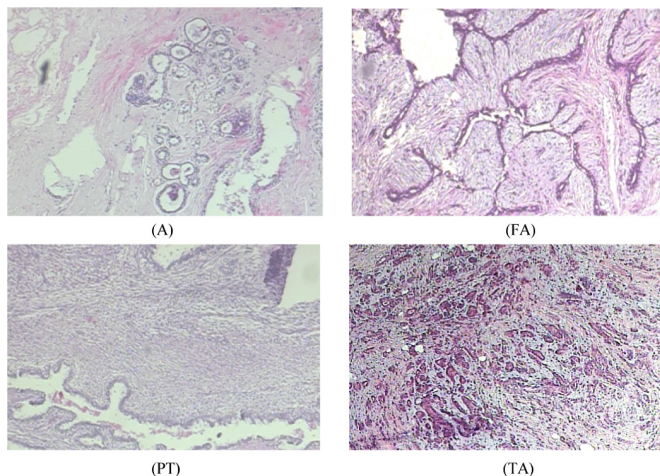


Fig. 1. Example of images of benign breast tumours stained with HE: A: Adenosis; F: Fibroadenoma; PT: Phyllodes Tumour; TA: Tubular Adenoma. Magnification factor = 400x.

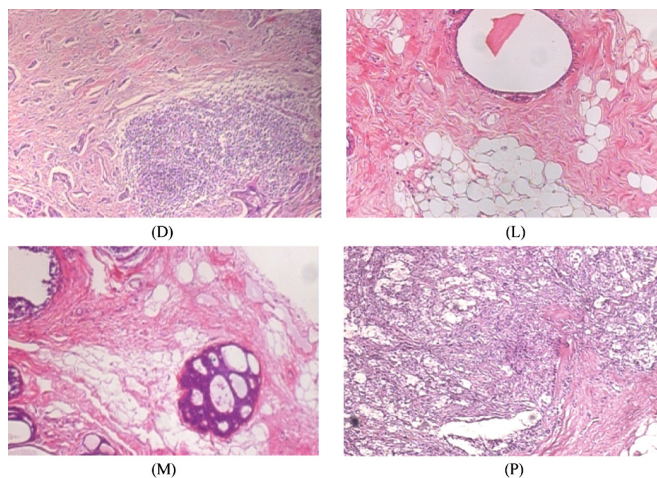


Fig. 2. Example of images of malignant breast tumours (breast cancer) stained with HE: D: Ductal; L: Lobular; M: Mucinous; P: Papillary. Magnification factor = 400x.

2.2 KAZE Based Bag of Features Classifier

KAZE is a novel 2D feature detection and description method that operates completely in a nonlinear scale space using additive operator splitting (AOS) algorithm and non-linear diffusion [24]. Nonlinear diffusion enables detecting and describing features in nonlinear scale spaces. Unlike SIFT [16] or SURF [17], that perform on Gaussian scale space, KAZE keeps important image information and removes noise.

During the training stage, KAZE keypoints were extracted for every image in the training set and used as feature vectors in a bag-of-feature approach. Bag-of-features technique is commonly used for image classification and is referred often as bag-of-words. After extracting KAZE features from grayscale images, a vocabulary was built reducing the number of features through quantization of feature space using a clustering algorithm. K-means algorithm was used to group features as localized clusters. The K-means algorithm iteratively grouped the feature descriptors into k mutually exclusive clusters. These clusters were separated by similar characteristics. Each cluster centre was a vocabulary word in a dictionary of visual words. Number of visual words to include in the bag-of-features is the vocabulary size and corresponds to K in the K-means clustering algorithm.

In this study, K was varied from 50 to 800 in order to explore the influence of vocabulary size in the classification accuracy.

The bag-of-visual words was used to encode images. Detected and extracted KAZE features were processed using the approximate nearest neighbour algorithm. The visual word occurrences in an image were counted producing a histogram that becomes a new and reduced representation of the image. The histogram length matches the number of visual words. The histogram encodes the image into a features vector, used as input to build (train) a binary support vector machines (SVM) classifier.

In the testing stage, the KAZE keypoints were extracted for every image in the test set, and fed into the cluster model to map them into a histogram vector. This vector was finally fed into the binary SVM classifier trained in the training stage, to classify the image. Figure 3 depicts the model applied in this work.

2.3 Evaluation Metric

Two main performance measurements were used to report accuracy of the developed model. The patient score for patient i (PS_i) was defined as:

$$PS_i = \frac{Irec_i}{Ip_i} \quad (1)$$

where Ip_i is the number of images for patient i and $Irec_i$ represents the number of images correctly classified for patient i .

The patient recognition rate (PRR) is given by:

$$PRR = \sum_{i=1}^N \frac{PS_i}{N} \quad (2)$$

where N is the total number of patients and PS_i is patient score of patient i .

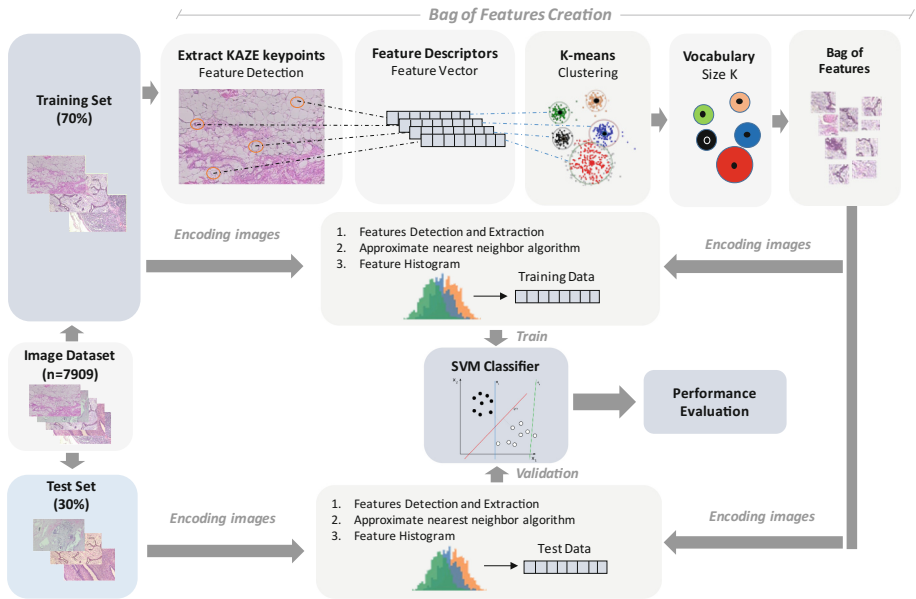


Fig. 3. Bag of Features model along the clustering and classification procedures.

Image recognition rate (*IRR*) was defined as the ratio of total correctly classified images (T_c) and the total number of images (T):

$$IRR = \frac{T_c}{T} \quad (3)$$

In addition to *PRR* and *IRR*, *F-score* was calculated for each case. *F-Measure* is the harmonic mean of precision (*Pr*) and recall (*Rc*) being both parameters defined as:

$$Pr = \frac{TP}{TP + FP} \quad Rc = \frac{TP}{TP + FN} \quad F = 2 \frac{Pr Rc}{Pr + Rc} \quad (4)$$

where *TP* is true positive, *FP* false positive and *FN* false negative cases. Sensitivity (*Se*) and specificity (*Sp*) were calculated as follows:

$$Se = \frac{TP}{TP + FN} \quad Sp = \frac{TN}{TN + FP} \quad (5)$$

Finally, area under the receiving operating characteristic (ROC) curve (AUC) was estimated for each model.

3 Results and Discussion

The BreakHis dataset was divided into a training (70%) and a testing (30%) subset. Patients who were used in the training set were not included in the testing set, to make sure the classifier generalizes to unseen patients. To enable the direct comparison of results with other state of the art methods, the same partitions for the five-fold replications used in [9] were applied in this study. Partitions are available to download¹.

KAZE features were extracted from all images and then encoded. Conductivity for KAZE features extraction was computed using regions.

The used number of scale levels was four and the selected multiscale detection factor was three. Vocabulary sizes of {50, 100, 200, 400 and 800} for the bag-of-features were evaluated.

C-SVM was used as model for the SVM classifier and radial basis function (RBF) was selected as kernel. A parameter optimization search control was applied using first a grid search and later, a pattern search over a narrow range around the best point found in the grid search.

Table 2 reports the accuracy of the proposed classification framework at both patient and image level, defined by Eqs. 2 and 3.

Table 2. Mean accuracy and standard deviation at image-level and patient-level of the SVM classifier in the test set, for different magnification factors and vocabulary sizes. Bold shows the best results over magnification factors.

	Magnification	Vocabulary size				
		50	100	200	400	800
Image level	40x	80.0 ± 2.7	82.6 ± 3.1	84.1 ± 2.5	85.0 ± 1.5	85.9 ± 1.6
	100x	75.7 ± 1.3	76.8 ± 1.4	78.3 ± 1.6	79.5 ± 2.0	80.4 ± 1.4
	200x	74.7 ± 0.9	76.0 ± 1.1	77.1 ± 1.4	77.8 ± 1.2	78.1 ± 2.2
	400x	69.6 ± 2.5	71.3 ± 1.8	71.1 ± 2.2	71.1 ± 2.1	71.1 ± 3.3
Patient level	40x	80.4 ± 2.5	82.6 ± 3.1	84.3 ± 2.2	85.3 ± 1.7	86.4 ± 2.2
	100x	77.2 ± 0.9	78.3 ± 1.5	79.5 ± 1.4	80.7 ± 2.3	81.6 ± 1.6
	200x	73.8 ± 1.5	75.3 ± 1.1	76.6 ± 2.1	77.5 ± 0.6	77.8 ± 1.6
	400x	70.5 ± 3.1	71.2 ± 2.0	72.2 ± 2.0	72.6 ± 1.9	72.9 ± 2.8

Table 3 reports the F_1 -scores at image level for each magnification factor and vocabulary size. Best results were obtained for a vocabulary size of 800. Tests with a greater vocabulary size were carried out but without significant improvement in accuracy.

¹ <https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/>.

Table 3. Mean F_1 -scores and standard deviation at image-level calculated in the test set over different magnification factors and vocabulary sizes. Bold shows the best results.

Magnification	Vocabulary size				
	50	100	200	400	800
40x	85.9 \pm 2.2	87.5 \pm 2.3	88.7 \pm 1.8	89.3 \pm 1.2	90.2 \pm 1.1
100x	83.3 \pm 0.9	84.1 \pm 0.9	85.1 \pm 1.1	85.7 \pm 1.6	86.5 \pm 1.0
200x	82.4 \pm 0.8	83.2 \pm 1.2	84.1 \pm 1.2	84.5 \pm 1.4	84.6 \pm 2.2
400x	78.8 \pm 2.1	80.3 \pm 1.4	80.1 \pm 2.1	79.9 \pm 2.1	79.9 \pm 3.2

Table 4 shows a comparison between approaches reported in different studies that used the BreaKHis database and results of this work, at both image and patient reporting level. As can be appreciated, the combination of KAZE keypoints and bag-of-features seems a suitable strategy for a low magnification factor. In the case of the 40x magnification factor, the proposed framework achieved the best image level accuracy with a mean recognition rate of 85.9% for a vocabulary size of 800. In such a case, achieved sensitivity was 95.9% while specificity was 73.7%. AUC was 0.94.

Table 4. Comparison of obtained results (vocabulary size = 800, mean accuracy) with existing literature using the BreaKHis database.

	Method	Magnification			
		40x	100x	200x	400x
Image level	Deep features [29]	84.6	84.8	84.2	81.6
	PFTAS [9]	82.8	80.7	84.2	81.2
	This work	85.9	80.4	78.1	71.1
Patient level	Deep learning [14]	83.0	83.1	84.6	82.1
	Deep features [29]	84.0	83.9	86.3	82.1
	PFTAS [9]	83.8	82.1	85.1	82.3
	This work	86.4	81.6	77.8	72.9

Furthermore, patient and image level accuracy of the proposed approach for a magnification factor of 40x was greater than the reported in [9, 14, 29] for any magnification factor. It confirms that 40x seems to be the most discriminating magnification factor for KAZE keypoints. For higher magnification factors, KAZE features did not improve results in terms of accuracy.

Mean F_1 score at image level for a vocabulary size of 800 and magnification factor of 40x was 90.2%, what outperforms scores reported at [9, 29] for the same factor.

Table 5 details the error distribution according to the vocabulary size for a magnification factor of 40x. More than 40% of the classifications errors were false positives associated to the classification of adenosis and fibroadenoma as malignant tumours. Fibroadenoma presents similar properties to a malignant tumour what may imply that new approaches would be needed for the discrimination of these specific cases [9, 11].

Table 5. Mean error distribution (%) and standard deviation for the proposed approach (magnification factor of 40x) in the test set over different subclasses and vocabulary sizes. A: Adenosis; F: Fibroadenoma; PT: Phyllodes Tumour; TA: Tubular Adenoma; D: Ductal; L: Lobular; M: Mucinous; P: Papillary.

		Vocabulary size				
		50	100	200	400	800
Benign	A	26.7 ± 5.5	28.7 ± 4.9	29.5 ± 6.1	30.6 ± 6.6	31.7 ± 6.8
	F	15.5 ± 5.8	13.3 ± 4.7	13.0 ± 3.6	13.1 ± 5.1	14.3 ± 6.2
	PT	5.6 ± 2.6	4.6 ± 2.0	4.3 ± 1.9	4.4 ± 3.0	4.9 ± 3.0
	TA	21.9 ± 5.0	19.9 ± 5.5	21.9 ± 4.8	21.4 ± 6.5	21.9 ± 8.5
Malignant	DC	12.6 ± 2.8	12.2 ± 3.5	10.3 ± 1.2	9.6 ± 2.8	9.1 ± 3.1
	LC	2.1 ± 1.8	2.2 ± 2.1	1.9 ± 1.4	2.1 ± 2.0	1.7 ± 1.8
	MC	10.1 ± 3.1	14.3 ± 3.9	12.2 ± 5.9	11.4 ± 1.7	9.1 ± 3.0
	PC	5.4 ± 4.5	4.9 ± 4.0	6.8 ± 4.2	7.3 ± 6.8	7.3 ± 6.2

4 Conclusions

In this paper, a classification framework based on the use of KAZE features and a bag of features classifier was presented and evaluated. The experimental results obtained on the BreakHis database confirm the feasibility of the proposed approach for the binary classification of breast cancer histopathological images at low magnification factors. Future work includes the exploration of different pre-processing techniques in order to improve the classification accuracy.

Acknowledgement. This work was supported by grant TIN2015-67020-P (Spanish “Ministerio de Economía y Competitividad and FEDER funds).

References

1. World Cancer Research Fund International: Cancer facts and figures - Worldwide data. <http://www.wcrf.org/int/cancer-facts-figures/worldwide-data>. Accessed 18 Jan 2018
2. Parks, R.M., Derks, M.G.M., Bastiaannet, E., Cheung, K.L.: Breast cancer epidemiology. In: Wyld, L., Markopoulos, C., Leidenius, M., Senkus-Konefka, E. (eds.) *Breast Cancer Management for Surgeons*, pp. 19–29. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-56673-3_3
3. Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J.W.W., Comber, H., Forman, D., Bray, F.: Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur. J. Cancer* **49**, 1374–1403 (2013)
4. World Health Organization: Breast cancer: prevention and control. <http://who.int/cancer/detection/breastcancer/en/index1.html>. Accessed 18 Jan 2018
5. Senkus, E., Kyriakides, S., Ohno, S., Penault-Llorca, F., Poortmans, P., Rutgers, E., Zackrisson, S., Cardoso, F.: Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **26**, v8–v30 (2015)
6. Veta, M., Pluim, J.P.W., Van Diest, P.J., Viergever, M.A.: Breast cancer histopathology image analysis: a review. *IEEE Trans. Biomed. Eng.* **61**, 1400–1411 (2014)

7. Aswathy, M.A., Jagannath, M.: Detection of breast cancer on digital histopathology images: present status and future possibilities. *Inform. Med. Unlocked* **8**, 74–79 (2017)
8. Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B.: Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* **2**, 147–171 (2009)
9. Spanhol, F.A., Oliveira, L.S., Petitjean, C., Heutte, L.: A dataset for breast cancer histopathological image classification. *IEEE Trans. Biomed. Eng.* **63**, 1455–1462 (2016)
10. Wang, P., Hu, X., Li, Y., Liu, Q., Zhu, X.: Automatic cell nuclei segmentation and classification of breast cancer histopathology images. *Sig. Process.* **122**, 1–13 (2016)
11. Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J., Monczak, R.: Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Comput. Biol. Med.* **43**, 1563–1572 (2013)
12. Zhang, Y., Zhang, B., Coenen, F., Lu, W.: Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Mach. Vis. Appl.* **24**, 1405–1420 (2013)
13. Zheng, Y., Jiang, Z., Shi, J., Ma, Y.: Retrieval of pathology image for breast cancer using PLSA model based on texture and pathological features. In: 2014 IEEE International Conference on Image Processing (ICIP), pp. 2304–2308. IEEE (2014)
14. Spanhol, F.A., Oliveira, L.S., Petitjean, C., Heutte, L.: Breast cancer histopathological image classification using convolutional neural networks. In: International Joint Conference on Neural Networks (IJCNN 2016), pp. 2560–2567 (2016)
15. Cruz-Roa, A., Basavanahally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A.: Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. Presented at the 20 March 2014
16. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**, 91–110 (2004)
17. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: Leonardi, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006). https://doi.org/10.1007/11744023_32
18. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2564–2571 (2011)
19. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 971–987 (2002)
20. Leutenegger, S., Chli, M., Siegwart, R.Y.: BRISK: binary robust invariant scalable keypoints. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2548–2555 (2011)
21. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, pp. 886–893 (2005)
22. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. *Image Vis. Comput.* **22**, 761–767 (2004)
23. Alahi, A., Ortiz, R., Vandergheynst, P.: FREAK: fast retina keypoint. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 510–517 (2012)
24. Alcantarilla, P.F., Bartoli, A., Davison, A.J.: KAZE features. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012*. LNCS, vol. 7577, pp. 214–227. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33783-3_16

25. Demchev, D., Volkov, V., Kazakov, E., Sandven, S.: Feature tracking for sea ice drift retrieval from SAR images. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 330–333. IEEE (2017)
26. Inoue, R., Goto, T., Hirano, S.: Authenticity inspection by image recognition using feature point matching. In: 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE), pp. 1–2. IEEE (2017)
27. Utaminingrum, F., Kumiawan, T.A., Fauzi, M.A., Wihandika, R.C., Adikara, P.P.: Adaptive human tracking for smart wheelchair. In: 5th International Symposium on Computational and Business Intelligence, ISCBI 2017, pp. 10–13 (2017)
28. Chen, Y.S., Chien, J.C., Lee, J.D.: KAZE-BOF-based large vehicles detection at night. In: 2016 International Conference on Communication Problem-Solving, ICCP 2016, pp. 2–3 (2016)
29. Spanhol, F.A., Cavalin, P.R., Oliveira, L.S., Petitjean, C., Heutte, L.: Deep features for breast cancer histopathological image classification. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 1868–1873 (2017)

Biomedical Signal Analysis



Low Data Fusion Framework Oriented to Information Quality for BCI Systems

Miguel Alberto Becerra^{1(✉)}, Karla C. Alvarez-Uribe¹,
and Diego Hernán Peluffo-Ordoñez²

¹ Instituto Tecnológico Metropolitano, Medellín, Colombia
migb2b@gmail.com

² Yachay Tech, Urcuquí, Ecuador

Abstract. The evaluation of the data/information fusion systems does not have standard quality criteria making the reuse and optimization of these systems a complex task. In this work, we propose a complete low data fusion (DF) framework based on the Joint Director of Laboratories (JDL) model, which considers contextual information alongside information quality (IQ) and performance evaluation system to optimize the DF process according to the user requirements. A set of IQ criteria was proposed by level. The model was tested with a brain-computer interface (BCI) system multi-environment to prove its functionality. The first level makes the selection and preprocessing of electroencephalographic signals. In level one feature extraction is carried out using discrete wavelet transform (DWT), nonlinear and linear statistical measures, and Fuzzy Rough Set – FRS algorithm for selecting the relevant features; finally, in the same level a classification process was conducted using support vector machine – SVM. A Fuzzy Inference system is used for controlling different processes based on the results given by an IQ evaluation system, which applies quality measures that can be weighted by the users of the system according to their requirements. Besides, the system is optimized based on the results given by the cuckoo search algorithm, which uses the IQ traceability for maximizing the IQ criteria according to user requirements. The test was carried out with different type and levels of noise applied to the signals. The results showed the capability and functionality of the model.

Keywords: Brain-computer interface · Data fusion · Evaluation system
Information quality

1 Introduction

The goal of data fusion (DF) is to reduce the amount of data and decrease their uncertainty. The DF impacts a lot of fields that use different types of data/information, and it is widely used by multiple applications such as brain computer interface (BCI) [1, 2], sensor networks, robotics, signal, image, and video processing, intelligent system design, among others [3, 4].

DF systems depend on the application and techniques selection, and their effectiveness is highly dependent on the IQ processed [5]. They cannot tackle all possible problems that affect the data (e.g., correlation, inconsistency, disparity among others)

[6]. Consequently, their developments are considered very complexly [7, 8]. Besides, they are sometimes ineffective to measure and improve the quality of information (IQ) since these measures depend on the context [9], and due to the diversity of the systems, multiple observations at different times, conditions, and experiments of the object or phenomenon which have multiple type of information [10]. In the same way, the performance of DF systems is not easy to measure, due to the difficulties in establishing general and complete indicators, the conflict among them, absence of generality, and limitations for applying in real environments [11].

In [12] a black box IQ methodology was proposed, that has been widely used as a reference. However, this does not allow tracing quality metrics into the process to make adequate adjustments to the DF process. There are multiple methodologies and frameworks for assessing the data/information quality (IQ) on information systems [13, 14], but few oriented to DF systems, therefore the reliability of assessing data or IQ of fusion systems is considered in [9] as a topic poorly explored and the most important problem, showing multiple challenges, including criteria and metrics definitions, quality control, data elimination, data modification, data presentation, among them, all based on IQ.

In [15] IQ and performance measures were proposed for DF systems applied to tracking. Then, in [16] is shown a general IQ assessment methodology for DF systems which considers the traceability together with global and local assessing based on a decomposition in elementary modules, which we consider as a complex task. In [17] was conducted a study of IQ in information fusion and decision making oriented to crisis management, highlighting the relevance of the context of each application for making the quality assessment. Her study is based on the Joint Directors of Laboratories (JDL) model [18], which is the most used DF model, and emulates the fusion that makes human beings from data and information obtained from their senses, and the cognitive processes for decision making support.

We selected the JDL model as reference for this work taking into account the problem of decomposing processes of a DF system but only was covered the level 0 (L0), and L1, and L4, which is known as low DF. This study proposed a model which aims at establishing specific criteria for IQ and performance by levels to obtain local (by levels) and global measures, mapping the qualities among levels to show the results to the user, and optimizing the fusion processes. This model tries to compile previously reported works in one model, following the four steps of the Total Data Quality Management (TDQM) cycle [19], where the contextual information can be integrated together with the quality requirements of the user in order to obtain a complete and general DF model that allows the reuse, integration, and modification. The model was validated with a BCI multi-environment based on a previous work discussed in [20], which studies the electroencephalographic (EEG) signals acquired from controlled, semi-controlled and uncontrolled environment for generating control signals. In our new study, we added different noise levels to EEG signals to identify general strengths, weaknesses and the capability of the system. The results demonstrated the functionality of our model and performance respect to typical BCI systems.

2 Background

2.1 Joint Directors of Laboratories (JDL) Model

The model is shown in Fig. 1 [18]. It has six levels as follows: (i) *L0- Source pre-processing* establishes the time, type and identity of the collected data; (ii) *L1-Object Refinement* carries out an iterative process of DF to track and identify targets and attributes of them. Therefore, it makes tasks on the data such as spatial-temporal alignment, association, *correlation*, tracking, estimation, clustering, and removing imperfect data, among others; (iii) *L2 - Situation refinement* establishes relationships among entities in order to generate an abstract interpretation based on inferences, (iv) *L3-Threat refinement* uses the relationships detected in level 2 for predicting and assessing possible threats, vulnerabilities, and opportunities in order to plan responses; (v) *L4-Process refinement* carries out the management of the sources tuning and adjusting. This process interacts with each level of the system; (vi) *L5-User refinement* delineates the human into the process of refinement. Besides, this system uses the Database Management, which controls the DF processes.

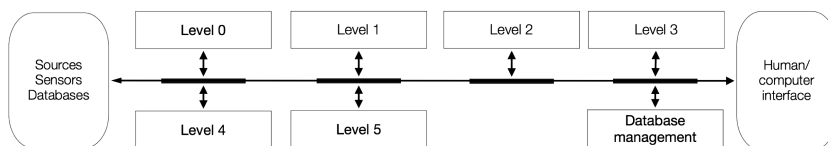


Fig. 1. JDL model

2.2 Information Quality Assessing

The information quality is an ambiguous term hard to describe and very subjective, and its evaluation is considered very complex, that must consider the requirements of the user since the interpretation of the quality depends on who will use the information [21]. Evaluating the IQ depends on different criteria, which have multiple metrics and their selection is based on user goals, process, and functions. However, assign quality attributes is a complex task in order to obtain one overall measure, therefore, one has to take into account the hierarchy of these and the user, context and process requirements which are consider by [9] as a similar problem of multi-criteria decision making which need comparing several alternatives based on values of the criteria under consideration. And depending on the application, cost can be considered for selecting the IQ measures [22].

For proposing a set of quality criteria, we followed the Total Data Quality Management-TDQM cycle methodology, which was performed in an iterative way. This has four components: (i) Define component identifies the needs of user and establishes IQ criteria. (ii) Measure component defines quality metrics and applies them. (iii) Analyze component identifies problems and impact of IQ measured, and (iv) improve component use different techniques for increasing the IQ [19].

Information Quality Criteria

The literature reports different studies of quality criteria together with their metrics used in information systems [12, 14, 22, 23], and DF systems [9, 16]. A criterion set was selected and analyzed to define its applicability on DF for levels 0 and 1 of the JDL model. Table 1 shows selected and grouped criteria taking into account the meaning (same or opposite-op), e.g., consistency and ambiguity have opposite meaning.

Table 1. IQ selected criteria

Criteria	Definition	Levels
Accuracy/ inaccuracy (op)	Intrinsic- validity of the information. Relational/contextual - Degree to which information correctly represents another information in the context of a particular activity or culture	0–1
Data amount	Define the availability of amount of data	0–1
Trust	Including believability, reliability, and reputation, catching how much information derives from an authoritative source	0–1
Credibility	Are regarded as true ad believable by users	0–1
Correctness	Correctly represent the true value of the intended attribute of a concept or event in a specific context of use	0–1
Cohesion/ consistency/ ambiguity (op)	The capability of the information to comply without contradictions to all properties of the reality of interest. Information is consistently represented using the same structure, format, and precision	0–1
Up to date/currency/ currentness	The age of the information and right age in a specific context of use	0–1
Minimality	Data that is not possible to eliminate some element from the schema without compromising the information content	0–1
Compactness/ conciseness/ redundancy (op)	It is referred to lack of redundancy (duplicates or same Information)	0–1
Relevance (aboutness)	Applicability of the information in a given activity	0–1
Efficiency	Can be processed and provide the expected levels of performance by using the appropriate amounts and types of resources in a specific context of use	0–1
Traceability	Provide an audit trail of access to the data and any changes made to the data in a specific context of use	0–1
Portability	Enable it to be installed, replaced or moved from one system to another preserving the present quality in a specific context of use. Traceability	Among levels
Completeness/precision/ incompleteness(op) repeatability/ reproducibility	The granularity or precision of the model or content values of information object according to some general purpose or needs in the context of a given activity. Non- empty elements	0,1
Reliability	It indicates “whether the data can be counted on to convey the right information; it can be viewed as correctness of data”. No interpretation regarding data deficiencies is given	0

Performance

Table 2 shows a set of performance metrics discussed in [11] from a literature review of performance of DF systems. Based on this work we established the number of citations of each performance (NC) as first criteria to set performance metrics for our model. The performance qualities selected for each level are shown in the next section.

Table 2. Performance metrics – NC

Metric	CF	Metric	CF	Metric	CF
ROC/AUC curve/correctness/false pos, false neg	4	Spatio-temporal info. preservation	1	Fusion/relative gain	2
Detection rate/false alarm/detection probability	7	Distance/position/performance/probability error	10	Signal to noise ratio	2
Response/execution/time	9	Credibility	1	Relevance	1
Accuracy	5	Reliability	1	Purity	1
Computation/cost	3	Mutual information	1	Confidence	1
Entropy	2				

3 Proposed Framework

Our proposed general low-level DF framework is depicted by L0 and L1, which includes (i) processing of hard data and contextual information, (ii) Assessing of IQ and performance (iii) optimization of the DF process.

The L0 is shown in Fig. 2, and it has the following functions: (i) Source selection is carried out from different sources S_1, S_2, \dots, S_n including contextual data SC_1, SC_2, \dots, SC_m ; (ii) Preprocessing: filtering, reconstruction of missing values, and normalization among others; (iii) Spatio-temporal alignment; and (iv) feature extraction: It extracts attributes of the objects. These functions are controlled by an expert system through rules, which are activated based on the results given by the local evaluation block, which assesses the IQ of the outputs and performance of the DF systems. The expert system, weighted mapping, and local evaluation blocks are included into the Level 4.

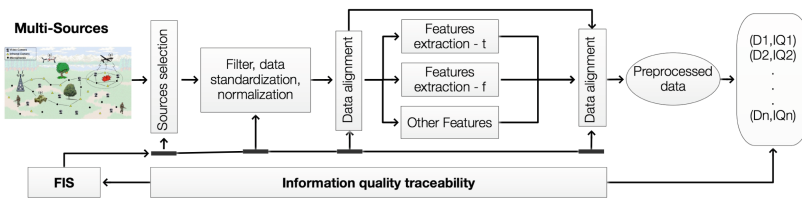


Fig. 2. Level 0 of the proposed data fusion framework

Figure 3 shows the L1 of our proposed framework. This stage has the following functions: (i) Data alignment: it carries out the spatio-temporal alignment of the data sources together with the features depending on the application. (ii) Dynamic feature selection carries out the feature/data selection to reduce the dimensionality of the data, and obtain the relevant features and data, besides; this stage can include techniques for data association; (iii) Identification and tracking. (iv) Spatio-temporal alignment is carried out again to align different identifications or detection given by different algorithms or functions. (v) Clustering and ensemble to reduce data and create groups from identified objects. These functionalities are controlled by the expert system in a similar manner to the L0.

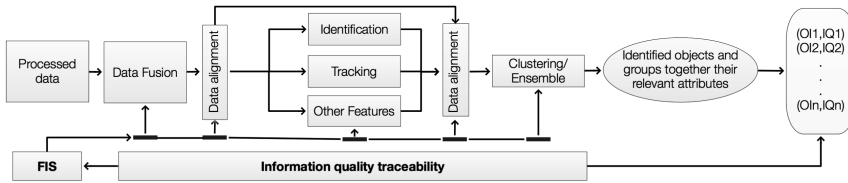


Fig. 3. Level 1 of the proposed data fusion framework

3.1 Information Quality and Performance Quality Characterization

This characterization was carried out by level taking into account the frequency of use reported in the literature and specific works showed in the last section trying to avoid redundancy. We established two types of quality criteria by level: IQ and quality oriented to the performance of the DF system. Table 3 shows the IQ criteria, and Table 4 shows the performance quality criteria that we selected for L0. Tables 5 and 6 show the IQ criteria, and performance quality criteria respectively for L1.

Table 3. Data quality criteria for Level 0 (output and input)

Data amount	Up to date/currency	Reliability
Correctness	Accuracy	Relevance
Consistency	Interpretability	Completeness

Table 4. Performance criteria for Level 0

Delay/timeliness	Computational cost	Robustness
Throughput	Sample rate	Accuracy

Table 5. Data quality criteria for Level 1 (output)

Repeatability	Data amount	Relevance
Consistency	Up to date/currency	Completeness

Table 6. Performance criteria for Level 1

Accuracy	Computational complexity	Throughput
Delay/timeliness	Accessibility	

3.2 Information Quality and Performance Evaluation

Figure 4 shows the IQ mapping system of the general DF model, which is used to present the quality traceability among levels to the user, and improve the results delivered by the DF process. This system has two levels (0 and 1) of the JDL model. The L4 and L5 of the JDL are immersed into the optimization process together with the user interaction respectively.

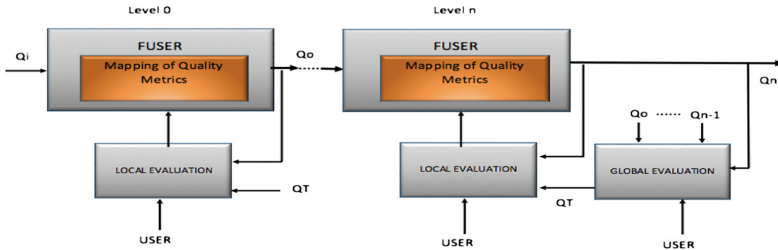


Fig. 4. Information quality system

In each level, a local IQ measure QM_{Ln} is obtained by mean of the equation $QM_{Ln} = f(Q_{n0}, Q_{n1}, \dots, Q_{nx})$, which is defined as $QM_{Ln} = \sum_{x=1}^r Q_{nx}/r$. Where, n is the level number, and x is the number of the output variable and r is the total number of variables of the level n . The quality measure of each variable is established based on different quality criteria weighted by the users. It is defined as $Q_{nx} = [w_{x1}q_{x1} + w_{x2}q_{x2} + \dots + w_{xc}q_{xc}]$. Where q_{xc} are the quality criteria of each variable v_{nx} , and w_{xc} are the weight values for each quality criteria, which are selected by the users and these are constrained to $\sum_1^c w_c = 1; 0 \leq w_c \leq 1, w_c$.

On the other hand, each criterion is defined by p quality metrics, which are selected by the user depending on the application or the environment. These quality metrics are expressed as $q_{xc} = [m_1 m_2 \dots m_p]$. Finally, the traceability is expressed by the following equation, where n is the level, x is the number of the datum, c is the criterion, and p is the quality metric.

$$IQT^{n \times c \times p} = \left\{ \begin{array}{ccc} \left[\begin{array}{ccc} m_{-111} & \dots & m_{-11p} \\ \vdots & \dots & \vdots \\ m_{-11c1} & \dots & m_{-11cp} \end{array} \right] & \dots & \left[\begin{array}{ccc} m_{-1x1} & \dots & m_{-1xp} \\ \vdots & \dots & \vdots \\ m_{-1xc1} & \dots & m_{-1xcp} \end{array} \right] \\ \vdots & \dots & \vdots \\ \left[\begin{array}{ccc} m_{n11} & \dots & m_{n1p} \\ \vdots & \dots & \vdots \\ m_{n1c1} & \dots & m_{n1cp} \end{array} \right] & \dots & \left[\begin{array}{ccc} m_{nx1} & \dots & m_{nxp} \\ \vdots & \dots & \vdots \\ m_{nxc1} & \dots & m_{nxcp} \end{array} \right] \end{array} \right\}$$

The Local Assessment

The local quality measure is defined as $ML_n = f(p_1QM_{nL} + p_2MOP_{nL})$, where, M_{Ln} : Total Local Quality Measure, QM_{nL} : *LocalIQ*, MOP : performance measure, where the weight values p_x for each measure are adjusted by the user according to requirements for each local measure; $\sum_1^n p_n = 1; 0 \leq p_n \leq 1$.

The Global Evaluation

The global measure is established adding the local measures as it is described by $M_G = f(a_{L0}ML_0 + a_{L1}ML_1)$, where, a_{Ln} : binary weight factor of Local Measure, which allows eliminating levels that are not used.

Weighted Mapping

IQ mapping is carried out taking into account the nonanalytic evaluation suggested by [16] which is considered adequate for a large number of criteria with dependencies among them and lack of process system knowledge. Then, the mapping is carried out using a machine learning-regression (MLR) $f(Q, \alpha)$ (see Fig. 5 – we applied Artificial Neural Networks (ANN)). The mapping of quality criteria among levels is described by $Q_{cnm} = f(Q_{c(n-1)1}, Q_{c(n-1)2}, \dots, Q_{c(n-1)x})$. Where Q_{cnm} is the local quality of the level n of the quality criteria m ; x defines the number of the quality criteria of the previous level.

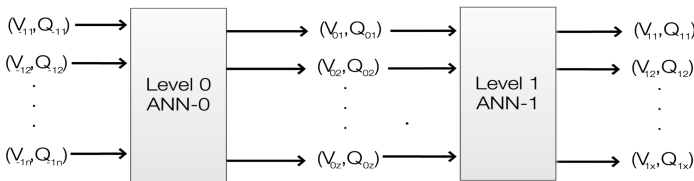


Fig. 5. Quality mapping system

Each level is trained using an observation set $\{(V_{c(n-1)1}, V_{c(n)x}), \dots, (V_{c(n-1)m}, V_{c(n)x})\}$ alongside their respective qualities $\{(Q_{c(n-1)1}, Q_{c(n)x}), \dots, (Q_{c(n-1)m}, Q_{c(n)x})\}$ where α is the parameter of the selected function that allows approximating the quality outputs $Q_{c(n)x}$ based on the quality observations $Q_{c(n-1)m}$. The trained MLR generates the estimated output set $\{\hat{Q}_{cn1}, \hat{Q}_{cn2}, \dots, \hat{Q}_{cnm}\}$ from the input set $\{Q_{c(n-1)1}, Q_{c(n-1)2}, \dots, Q_{c(n-1)z}\}$. A general regression system is depicted as $\hat{Q}_{cny} = \sum_{i=1}^m \beta_i(x)Q_{c(n-1)i}$.

The functionality of this mapping system is executed following three steps: (i) If there not exist a historical data set, is necessary to generate them applying synthetic data with different noise levels and build a quality data set. (ii) The MLR is trained using the data set obtained in the previous step. (iii) The mapping system for each input set, it gives the IQ by level together with the IQ relation among the different levels. (iv) The mapping system is retrained with each newly executed process.

This system defines the relation among quality of inputs and quality of outputs, and it is depicted with two stages as follows: (i) Stage 1 is the mapping of quality metrics, which supports the traceability among levels. (ii) Stage 2 establishes the dependence

among quality criteria of inputs and quality criteria of outputs using calculated weights, maximizing the quality outputs, and applying one optimization algorithm to the mapping model.

User Functionality

Figure 6 shows the interaction between the user and IQ evaluation system. First, the system shows the quality criteria set with specific weights based on analytics applied to user’s preferences historical dataset (recommender system). Then, the user can adjust these weights according to his requirements. The number of criteria is different for each level, and the criteria were discussed in the previous sections. The evaluation results can be shown as a unified global measure M_G , or a set of local unified measures (one by level) $[QM_{L-1}QM_{L0}QM_{L1}\dots QM_{L3}]$. In order to make traceability of the quality among levels, a matrix of quality criteria by level $Q_{nm \times x}$. This matrix is detailed to a matrix $IQT^{nm \times xp}$, which include the quality criteria together their metrics for each datum.

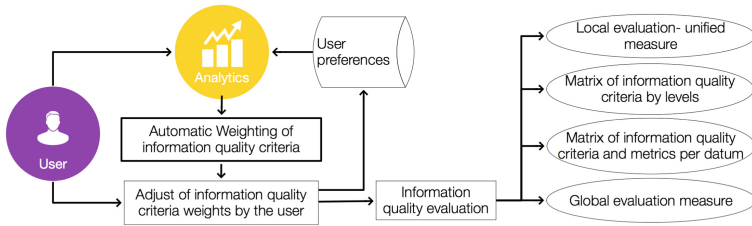


Fig. 6. User functionality

Optimization Stage

A multi-objective optimization algorithm is applied using the functions of traceability achieve with the traceability systems to establish the quality criteria that must improve them.

Expert System

We proposed the optimization by level of the DF process taking as objective ML_n and M_G on which the user can carry out some adjustments. These results together weights calculated, and contextual variables are used by the expert systems for making decisions on the DF systems to improve the quality of the inputs variables for obtaining an increase in the quality of output variables according to the relevance selected by the user. It is described through rules using a Fuzzy Inference System (FIS). The rules are defined as *if $QM_{Ln} < > alto$ and $Q_{cn1} = low$ then control action_{1,...}, if $QM_{Ln} < > alto$ and $Q_{cnm} = low$ then control action_m*. The control actions depend on the quality criteria, which are established by the optimization algorithm. For instance, if the quality criterion to adjust corresponds to the input of level 0 the selected data are applied depending on their quality value based on the threshold. Experts define these actions.

3.3 Validation

The proposed model was validated in a BCI multi-environment based on a previous work [20], where a BCI database was collected of 2 imaginary movements from 3 subjects in 4 different environment conditions. 14 EEG signals were acquired alongside 3 external variables (temperature, noise level (White Gaussian and thermal noise), drowsiness, concentration level, light level, and humidity. Additionally, each scenario or environment was assessed as rural, controlled, and urban. The same way each subject was assessed as external agent based on contextual information. Then, the proposed model was constructed with the following algorithms: (i) Butterworth filter and the Hampel algorithm were applied for removing noise and outliers respectively; (ii) the combination of the signals was carried out applying a Principal component analysis (PCA) algorithm; (iii) the feature extraction was carried out applying discrete wavelet transform (DWT), nonlinear and linear statistical measures, and Fuzzy Rough Set –RS algorithm for selecting the relevant features; (iv) The classification was carried out using support vector machine-SVM algorithm; (v) the expert system was constructed by mean the FIS; (vi) the cuckoo search algorithm was applied in order to optimized the objective function obtained from the mapping process; (vii) Mapping system was applied using Artificial Neural Networks. Finally, the users selected and weighted the quality measures. This proposed model was compared respect to a typical BCI system which is shown Fig. 7 (feature extraction (DWT, statistical measures + FRS) + SVM). The performance of the system was established using the accuracy measure (Table 7).



Fig. 7. Information quality mapping system

Table 7. Results proposed model respect to classical model

Noise level	Accuracy - classical system [%]	Accuracy - proposed model [%]
0 dB	94.6%	92.36%
10 dB	87.1%	90.7%
20 dB	72.9%	89.9%

4 Conclusions

In this study, we propose a complete DF model oriented to quality. We define a framework for assessing the IQ to optimize of the DF system across two levels of JDL DF model. Unlike others studies, we make a distinction among quality criteria by levels proposing specific criteria to establish one local measure by level and one global

measure, which are used for optimizing the fusion by level and the global fusion. Besides, contextual information was added and assessed. The model can be used to build any DF system, and its performance can be compared using any criterion proposed for each level. This allows to improve the development of these systems and to take advantage of the reuse of the developed systems in different application in an easier way considering the relevant quality criteria. The results demonstrated the capability of our model in respect to the classical processing systems in the BCI environment. In future researches, we will continue with the development of this model including the high fusion levels, and we will employ it in multiple real applications such as bioinformatics, biomedicine, Internet of things among others.

Acknowledgments. This work was supported by the Doctoral thesis “Data fusion model oriented to information quality” at the “Universidad Nacional of Colombia”.


References

1. Leeb, R., Sagha, H., Chavarriga, R., del R Millán, J.: A hybrid brain–computer interface based on the fusion of electroencephalographic and electromyographic activities. *J. Neural Eng.* **8**(2), 025011 (2011)
2. Oken, B., Orhan, U., Roark, B., Erdogmus, D., Fowler, A., Mooney, A., Peters, B., Miller, M., Fried-Oken, M.B.: Brain-computer interface with language model-electroencephalography fusion for locked-in syndrome. *Neurorehabil. Neural Repair* **28**(4), 387–394 (2014)
3. Smith, D., Singh, S.: Approaches to multisensor data fusion in target tracking: a survey. *IEEE Trans. Knowl. Data Eng.* **18**(12), 1696–1710 (2006)
4. Ardeshir Goshtasby, A., Nikolov, S.: Image fusion: advances in the state of the art. *Inf. Fusion* **8**(2), 114–118 (2007)
5. Hadzagic, M., Valin, P., Shahbazian, E.: Reliability and relevance in the Thresholded Dempster-Shafer Algorithm for ESM data fusion. In: *Information Fusion (FUSION)*, pp. 615–620 (2012)
6. Khaleghi, B., Khamis, A., Karray, F.O., Razavi, S.N.: Multisensor data fusion: a review of the state-of-the-art. *Inf. Fusion* **14**(1), 28–44 (2013)
7. Esteban, J., Starr, A., Willetts, R., Hannah, P., Bryanston-Cross, P.: A review of data fusion models and architectures: towards engineering guidelines. *Neural Comput. Appl.* **14**(4), 273–281 (2005)
8. Sidek, O., Quadri, S.A.: A review of data fusion models and systems. *Int. J. Image Data Fusion* **3**(1), 3–21 (2012)
9. Rogova, G.L., Bosse, E.: Information quality in information fusion. In: *2010 13th International Conference on Information Fusion*, pp. 1–8 (2010)
10. Lahat, D., Adaly, T., Jutten, C.: Challenges in multimodal data fusion. In: *2014 Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*, pp. 101–105 (2014)
11. van Laere, J.: Challenges for IF performance evaluation in practice. In: *2009 12th International Conference on Information Fusion, FUSION 2009*, pp. 866–873 (2009)
12. Wang, R.Y., Strong, D.M.: Beyond accuracy: what data quality means to data consumers. *J. Manag. Inf. Syst.* **12**(4), 5–33 (1996)
13. Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y.: AIMQ: a methodology for information quality assessment. *Inf. Manag.* **40**(2), 133–146 (2002)

14. Stvilia, B., Gasser, L., Twidale, M.B., Smith, L.C.: A framework for information quality assessment. *J. Am. Soc. Inf. Sci. Technol.* **58**(12), 1720–1733 (2007)
15. Blasch, E., Valin, P., Bosse, E.: Measures of effectiveness for high-level fusion. In: 2010 13th International Conference on Information Fusion, pp. 1–8 (2010)
16. Todoran, I., Lecornu, L., Khenchaf, A., Le Caillec, J.-M.: A methodology to evaluate important dimensions of information. *ACM J. Data Inf. Qual.* **6**(2–3), 23 (2015). Article no. 11
17. Rogova, G.L.: Information quality in information fusion and decision making with applications to crisis management. In: Rogova, G., Scott, P. (eds.) *Fusion Methodologies in Crisis Management*, pp. 65–86. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-22527-2_4
18. Steinberg, A.N., Bowman, C.L., White, F.E.: *Revisions to the JDL Data Fusion Lex. by JDL* (1991)
19. Wang, Y.R., Ziad, M., Lee, Y.W.: *Data Quality*. Kluwer Academic Publishers, Dordrecht (2002)
20. Ortega-Adarme, M., Moreno-Revelo, M., Peluffo-Ordoñez, D.H., Marín Castrillon, D., Castro-Ospina, A.E., Becerra, M.A.: Analysis of motor imaginary BCI within multi-environment scenarios using a mixture of classifiers. In: Solano, A., Ordoñez, H. (eds.) *CCC 2017. CCIS*, vol. 735, pp. 511–523. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66562-7_37
21. Mendes, P.N., Mühleisen, H., Bizer, C.: Sieve: linked data quality assessment and fusion, pp. 116–123 (2012)
22. Haug, A., Haug, A., Zachariassen, F., van Liempd, D.: The costs of poor data quality. *J. Ind. Eng. Manag.* **4**(2), 168–193 (2011)
23. Cabitza, F., Batini, C.: Information quality in healthcare. *Data and Information Quality. DSA*, pp. 403–419. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-24106-7_13



New Parameter Available in Phonocardiogram for Blood Pressure Estimation

Omari Tahar¹ , Ouacif Nadia¹, Benali Redouane², Dib Nabil²,
and Bereksi-Reguig Fethi²

¹ Department of Biomedical Engineering, Boumerdes University, Boumerdes, Algeria
omari.tahar@univ-boumerdes.dz

² Department of Biomedical Engineering, Tlemcen University, Tlemcen, Algeria

Abstract. Continuous and non-invasive measurement of blood pressure (BP) is of great importance. To achieve continuous and cuffless BP monitoring, pulse transit time (PTT) has been reported as a potential parameter. Nevertheless, this approach remains very sensitive, cumbersome and disagreeable for patients. This study proposes a new parameter available in phonocardiogram (PCG) to measure blood pressure. The PCG is processed and analyzed in order to measure the systolic and diastolic durations in order to study their correlation with an estimated PTT. The proposed approach evaluated on a developed data base of 37 subjects shows linear regression of two classes in distribution of diastolic duration with PTT estimated. Where, a good correlation coefficient is found for each class (class1: $R = 0.87$, class2: $R = 0.85$). This technic has significant potential to develop a new approach to measure blood pressure using PCG signal.

Keywords: Blood pressure · Pulse transit time · PPG signal
PCG signal · Diastolic duration

1 Introduction

Blood Pressure (BP) including systolic blood pressure (SBP) and diastolic blood pressure (DBP) is an important and vital sign of health care. SBP is the peak pressure in the arteries when the blood flows from the ventricles to the arteries during ventricular systole, and DBP is the minimum pressure in the arteries during ventricular diastole when the ventricles are full of blood. The metrical unit of blood pressure is called millimeters of mercury (mmHg). BP can be measured both invasively and non-invasively, using or not using cuff, non-continuously and continuously. Instantaneous BP cannot show real-time physiological status momentarily, while the continuous BP can provide superior data for predicting risk of serious cardiovascular events and making antihypertensive therapy decisions [5]. Therefore, the continuous, cuffless and non-invasive blood pressure estimation based on the pulse transit time (PTT) has gained increasing attention in recent years [9].

The PTT is the time taken by a one pulse wave to travel a distance between two site in the body. Usually the heart is considered us starting point and the index is the arrived point. To achieve this measurement, two systems are needed: electrocardiograph (ECG) and photoplethysmograph (PPG). In ECG signal, the peak R indicates the starting point, in which the ventricles are contracted and the blood is ejected out. In PPG signal, the peak of this signal (P) indicantes the arriving point. The timing difference between the two peaks determines the pulse transit time PTT as presented in Fig. 1. The Blood pressure is then estimated by using some fluid dynamics laws [3]. However, its use remains limited because it requires two different instruments with many sensors (three ECG electrodes and PPG fingertips optical sensors) which are completely disagreeable for the patient.

The necessity to find an alternative solution is of great importance. Among alternative solutions, the Phonocardiographic signal (PCG) gain great interests. This signal represents the recording of the mechanical activity of heart. Generally, through a simple stethoscope two significant sounds can be distinguished (S1 and S2) produced respectively by the closure of the atrio-ventricular valves and sigmoid [4]. The recording and processing of this signal can give more information about the heart and the physical phenomena on it. In the PCG trace (Fig. 1) the fluctuations S1 and S2 indicate respectively the first and second heart sound. The spectral characteristics of these sounds are related directly to the state of the valves and the hemodynamic of blood. The interval S12 indicates the systolic duration, it is the time taken by the ventricle to expulse out the blood. The interval S21 is the diastolic duration, and it is the time taken by the ventricle to be full again. The variation in cardiovascular hemodynamic is reflected directly on various physiological signals, including PCG signal. Therefore, it is sufficient to find an adequate parameter in this signal which allows a correct estimation of blood pressure.

In the literature, several attempts have been made to achieve this goal. The first was in 1965 by Sakamoto et al. [6], where the dependence of the systolic blood pressure (SBP) with the intensity of the first sound S1 is studied. The results show a good correlation between the peak systolic pressure of left ventricle and the intensity of S1. The analysis of the second cardiac sound S2 was also the subject of several studies. Effectively, this sound is generally accentuated in hypertensive patients, which developed therefore the hypothesis of a possible correlation between blood pressure and the spectral characteristics of S2 [1]. In 2006, Zhang and Zhang [10] present a study on the relationship between the SBP and the time RS2 delimited by the peak R of the ECG signal at the pick of the second sound S2 of PCG signal (see Fig. 1). The correlation between the two parameters was acceptable with a rate of $R=0.85$. Recently, in 2015, Shukla et al. [7] propose a new approach to measuring blood pressure using vascular transit time (VTT), defined as the difference in time between S1 and the peak of PPG signal. However, the database used in this study is limited to validate this approach.

This paper is an attempt to explore the correlation between PTT and the both systolic (S12) and diastolic (S21) durations. We started with the hypothesis that if the physiological signals (PCG, ECG and PPG) are correlated together and if PTT (obtained from ECG and PPG) is correlated with systolic blood pressure, then, certainly there is some correlation between PCG signal and blood pressure. Also, if RS2 and PTT are correlated with systolic pressure, then, RS2 (which is systolic duration) should be correlated with PTT. The methodology followed to prove this hypothesis is presented in below.

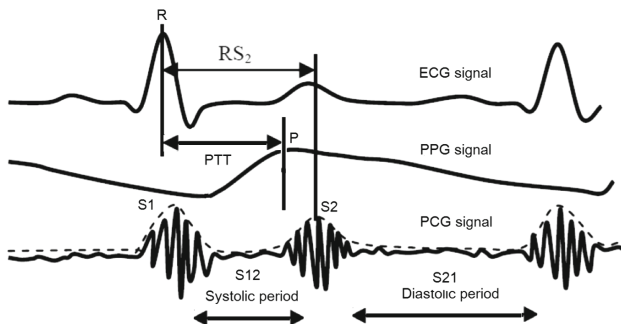


Fig. 1. Simultaneous presentation of the three physiological signals ECG PPG and PCG and the different parameter used in the estimation of the arterial pressure.

2 Methodology

To prove our hypothesis a dataset is done. This last contains the recording of PCG signal and the both systolic and diastolic blood pressure measured by sphygmomanometer for a number of volunteer. From PCG signal, the duration systolic (S12) and diastolic (S21) can be extracted. While, from the measured blood pressures the PTT of each subject can be estimated.

Before any collection of data, including biomedical data, three important points must be considered: the population studied, the protocol used in the measurement and the system of data acquisition. In the acquisition of physiological data, its necessary to determine a protocol of measurement in order to create the same conditions for all subjects.

The population studied in our work consists of 37 subjects, of different sexes and ages. It is constituted by a majority of adult (27 male and 10 female) with a range age of 22 to 40 years. The majority of the population is in good health, except for a few cases with a mild hypertension. The acquisition system used in this work is PowerLab 15 T made by ADInstrument. It is a powerful system in acquisition and processing of physiological signals in real time. This system can be connected to an electronic stethoscope of the same Firm (MTL209) with jack plug. The PCG signal is acquired with a sampling frequency of 10 khz and filtered by a lowpass filter with a cut-off frequency of 500 Hz. The Blood pressure

is measured by an automatic manometer made by Microlife, which allows direct measurement of the systolic and diastolic pressures and heart rate.

In order to put all subjects in the same measuring conditions, a measurement protocol has been established. It consists to take the measurements in sitting position after a resting period of 10 min. This rest time is necessary; it allows the stability of pressure in the body. After the rest time, the PCG signal is recorded for a 5 cycles then the blood pressure is measured immediately. The measure of blood pressure is repeated for three times then an average is determined. We note that the measurements of blood pressure were made in left arm. The pressure in this arm is more accurate to estimate the pressure in the heart.

The measurement of the systolic and the diastolic durations are made by the segmentation algorithm described in [2], the result of this operation is shown in Fig. 2.

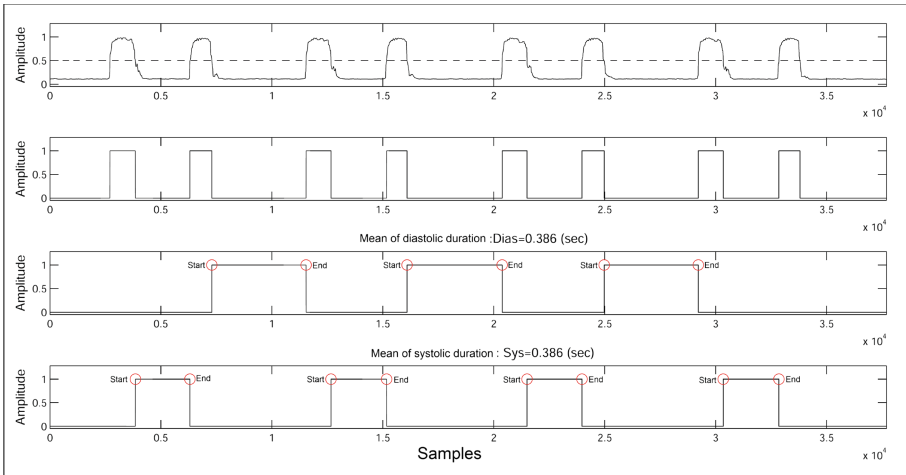


Fig. 2. The result of the different steps of the segmentation algorithm. (a): The detected envelop of PCG signal, (b): The maximisation of the envelop between [0,1] after thresholding operation, (c): The extracted diastolic phases with the mean duration value and (d): The extracted systolic phases with the mean duration value.

The PTT can be estimated using the physical laws in the body. Indeed, the work done by the pulse wave can be expressed in terms of the kinetic energy of the wave and the gravitational potential energy as shown in Eq. 1

$$F \cdot d = \frac{1}{2}mv^2 + mgh \tag{1}$$

where: F is the force exerted on blood, d is the distance from heart to wrist (or fingertip), m is the mass of blood, v is the pulse wave velocity, $g = 9.81 \text{ m/s}^2$, and h is the height difference between heart and wrist (or fingertip).

The force exerted on blood can also be written in terms of pressure difference as shown in Equation

$$F = \Delta BP \cdot a \tag{2}$$

where: a is the cross section area of the artery, and ΔBP is the pressure difference between the two sites, the heart and the wrist or fingertip.

Substituting Eq. 2 into Eq. 1 and rearranging the variables...

$$\Delta BP = \frac{1}{2} \frac{m}{a \cdot d} v^2 + \frac{m}{a \cdot d} gh \tag{3}$$

Replacing $m/(a.d)$ by ρ (density of blood) and v by d/PTT (pulse wave velocity of blood), Eq. 3 can be expressed as shown in Eq. 4.

$$\Delta BP = \frac{1}{2} \rho \frac{d^2}{PTT^2} + \rho gh \tag{4}$$

The distance d can be approximated by measuring the distance between the heart and the wrist as illustrated in Fig. 3. While, the PTT is the pulse transit time measured in seconds. Furthermore, the average blood density is reported at 1050 kg/m^3 [14th]. On the other hand, the pressure drop in the arterial side of the circulation is about 70% of the total pressure drop in the body [15th]. Therefore the subject's average change in the blood pressure or mean blood pressure, MBP, can be approximated using Eq. 5.

$$MBP = \frac{\Delta PB}{0.7} \tag{5}$$

In this case, the MBP unit obtained is in Pascal, for this purpose a change of unity towards the mmHg is necessary. This can be done by the multiplication of the value found by 0.0075.

$$MBP(mmHg) = \left(\frac{\rho}{1.4} \frac{d^2}{PTT^2} + \frac{\rho gh}{0.7} \right) 0.0075 \tag{6}$$

According to anthropometric segment length of human body [8], the distance d can be estimated by $(0.6 \times \text{height})$, similarly the distance h can be estimated in sitting position by $(0.30 \times \text{height})$. Finally, by replacing the constants (ρ and g) by their value, the MBP can be represented by two variables (PTT and height) as shown in Eq. 7. Which can be reformulated to get PTT dependence with MBP as given in Eq. 8. Where, MBP is the Mean blood pressure (mmHg), PTT is the Pulse transit time (sec).

$$MBP(mmHg) = \frac{1.947 \text{ height}^2}{PTT^2} + 31.84 \text{ height} \tag{7}$$

$$PTT(sec) = \sqrt{\frac{1.947 \text{ height}^2}{MBP - 31.84 \text{ height}}} \tag{8}$$

Where: PTT is the pulse transit time (sec), MBP is the mean blood pressure (mmHg) and height is the subject height(m).

In our study, the mean pressure (MBP) is calculated by Eq. (9). Where, DBP and SBP are diastolic and systolic blood pressure measured by manometer.

$$MBP(mmHg) = \frac{2}{3}DBP + \frac{1}{3}SBP \tag{9}$$

So, the PTT is determined for each subject of the database by measuring height and MBP of Eq.9. In the result section, the distribution of the both systolic and diastolic durations with the estimated PTT is discussed.

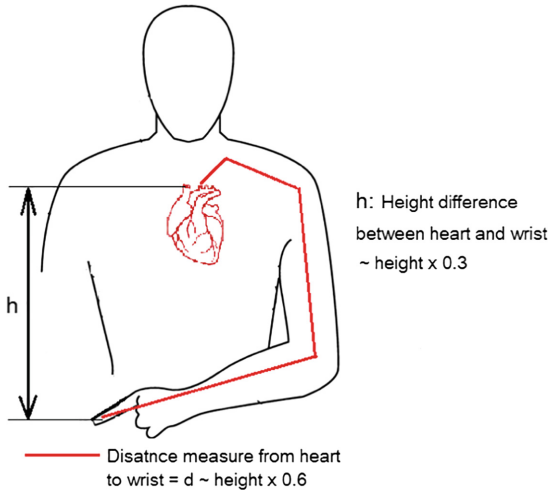


Fig. 3. Illustrations of distance measurements from heart to wrist.

3 Results

The distribution results of the different measures are presented in Fig. 4. In the Fig. 3(a), we can observe the distribution of systolic duration with PTT. From this last, none particular information can be extracted. All data are dispatched such as a cloud of point. In Fig. 3(b), the distribution of diastolic duration with PTT is presented. In this last we can show a particular distribution, where, a linear evolution of two classes can be clearly observed (see Fig. 5). These classes, can be approximated by two linear curves with a high correlation coefficient $R = 0.85$ for class:1 and $R = 0.87$ for class:2.

A comparison between the two classes according to the height of subject, the weight of subject, the systolic duration and the diastolic duration is done in order to find a probable parameter to establish the separating rule for these classes.

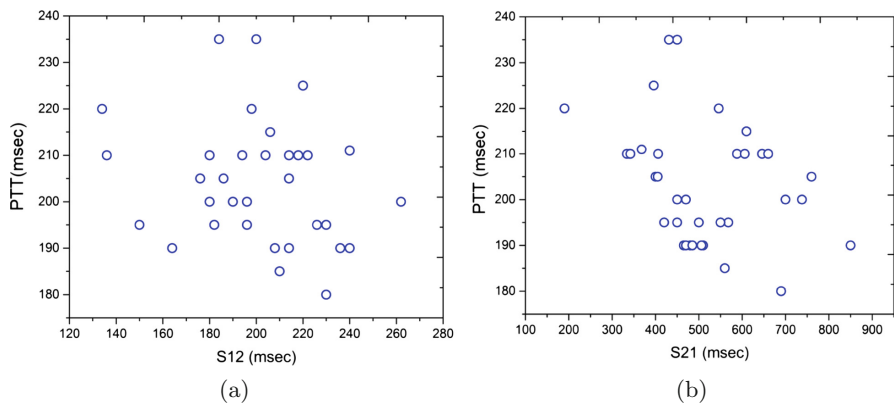


Fig. 4. The distribution of diastolic and systolic durations according to pulses transit time PTT. (a) the systolic duration (S12), (b): the diastolic duration (S21)

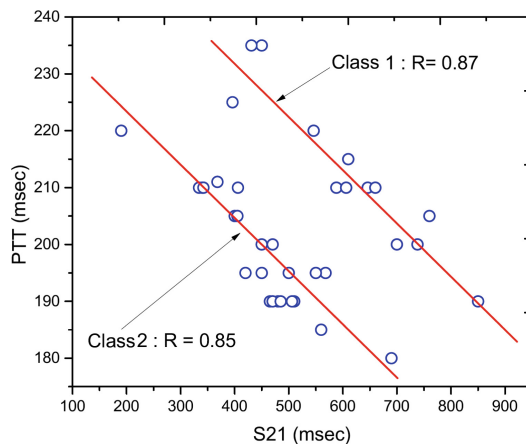


Fig. 5. The distribution of diastolic duration with the estimated PTT, and the approximation curve for each class.

The result of this study is shown as boxplot in Fig. 6. The comparison between the two classes according the height of subject is presented in Fig. 6(a). We can observe that the population of class 1 varies from 175 cm up to 187 cm, in which 50% of the population (bounded by the box) has a height in the range of 176 cm up to 182 cm. However, the extent of class 2 varies from 157 cm up to 185 cm, whose the height of 50% of the population of this class varies from 166 cm up to 178 cm. Although the range of class 2 is significantly larger than class 1, the interference between the two classes exists particularly from 175 cm up to 180 cm.

Figure 6(b) presents the comparison between the two classes according the weight of subject. We can observe that the weight of class1 population varies interval from 50 kg up to 96 kg, in which 50% of the population have a weight

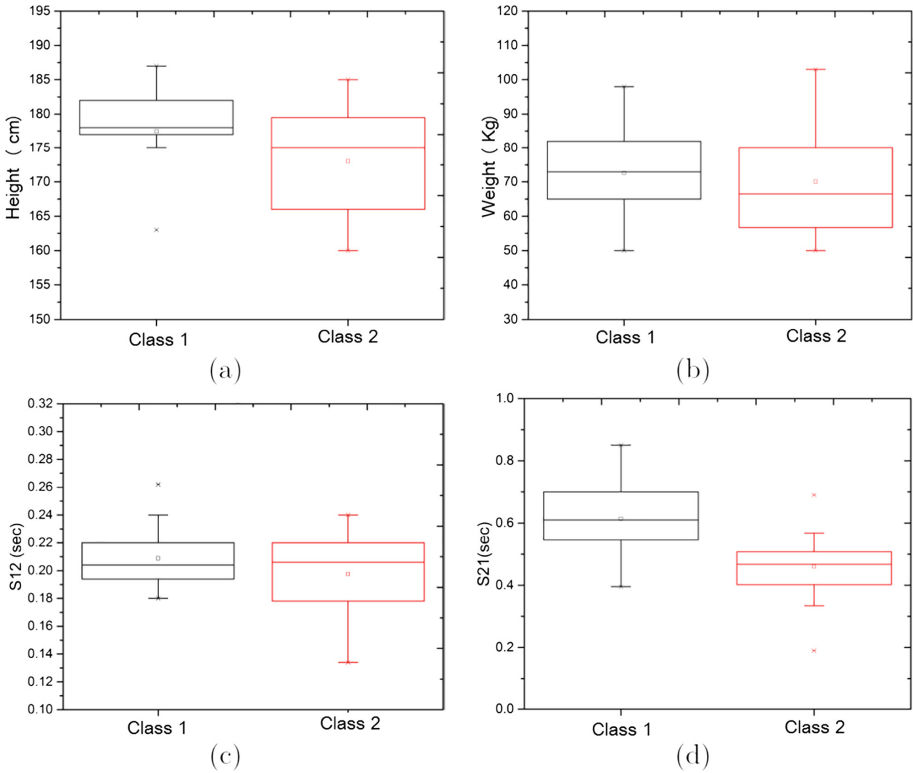


Fig. 6. The comparison between the two classes. (a): according to height, (b): according to weight, (c): according to systolic duration and (d): according to diastolic duration.

between 82 kg up to 82 kg. However, the weight of the population of class 2 varies between 78 kg up to 78 kg. Also, Interference between the two classes is existed in almost part of the distribution.

Figure 6(c) presents the comparison between the two classes according the systolic duration (S12). In the latter, we observe that the range of class 1 varies from 0.18 up to 0.24 s, in which the systolic duration of 50% of population is between 0.19 to 0.22 s. However, the extend of class 2 varies from 0.13 s up to 0.24 s, in which the systolic duration of 50% of population is between 0.17 to 0.22 s. Nevertheless, we can well notice the interference between the two classes in most parts of the distribution.

Figure 6(d) presents the comparison between the two classes according the diastolic duration (S21). We can observe that the S21 of class 1 varies from 0.4 s to 0.82 s, in which the diastolic duration of 50% of the population is between 0.5 to 0.7 s. While the diastolic duration of the population of class 2 varies in the interval ranging from 0.3 to 0.57 s, in which 50% of the population has a diastolic duration in the interval from 0.4 to 0.5 s. In this parameter, a slight separation is observed between the two boxes. However, the interference still exists between them especially in the range from 0.4 to 0.5 s.

4 Conclusion

In this paper, the correlation between the pulse transit time (PTT) and both systolic and diastolic duration gated from PCG signal is done. In this aim, a database composed by the recording of PCG signal and the measurement of blood pressures (systolic and diastolic) is done on 37 subjects. Also, an algorithm is developed to measure automatically the systolic (S12) and diastolic (S21) durations.

The statistical study shows a particular distribution of diastolic duration (S21) with PTT. Where, a linear regression of two classes is observed with a high coefficient of correlation (class:1 $R=0.85$, class:2 $R=0.87$). This result is very important; it allows a probable possibility to estimate blood pressure directly by diastolic duration, and particularly by PCG signal. In order to find a probable parameter able to separate each class, a comparison according to the height, the weight, the systolic and diastolic duration is done between the two classes. Sadly none seem able to make a significant separation.

In future works, we look to find the optimal parameter able to distinct each class by testing other variables. We look also to develop an intelligent system able to distinct these classes.

In summary, this study can be used to observe the dynamic changes in BP continuously using an electronic stethoscope. This technology has significant potential to advance healthcare, It can be incorporated easy in smartphone technologies.

Acknowledgement. This research was partially supported by Boumerdes university-Algeria. We thank our colleagues from both Tlemcen university and boumerdes university who provided insight and expertise that greatly assisted the research, although they may not agree with all of the conclusions of this paper.

References

1. Bartels, A., Harder, D.: Non-invasive determination of systolic blood pressure by heart sound pattern analysis. *Clin. Phys. Physiol. Meas.* **13**(3), 249 (1992)
2. Nigam, V., Priemer, R.: Simplicity based gating of heart sounds. In: 2005 48th Midwest Symposium on Circuits and Systems, pp. 1298–1301. IEEE (2005)
3. Obrist, P.A., Light, K.C., McCubbin, J.A., Hutcheson, J.S., Hoffer, J.L.: Pulse transit time: relationship to blood pressure. *Behav. Res. Methods* **10**(5), 623–626 (1978)
4. Omari, T., Bereksi-Reguig, F.: An automatic wavelet denoising scheme for heart sounds. *Int. J. Wavelets Multiresolut. Inf. Process.* **13**(03), 1550016 (2015)
5. Perloff, D., Sokolow, M., Cowan, R.: The prognostic value of ambulatory blood pressures. *JAMA* **249**(20), 2792–2798 (1983)
6. Sakamoto, T., Kusakawa, R., Maccanon, D.M., Luisada, A.A., Harvey, I.: Hemodynamic determinants of the amplitude of the first heart sound. *Circ. Res.* **16**(1), 45–57 (1965)

7. Shukla, S.N., Kakwani, K., Patra, A., Lahkar, B.K., Gupta, V.K., Jayakrishna, A., Vashisht, P., Sreekanth, I.: Noninvasive cuffless blood pressure measurement by vascular transit time. In: 2015 28th International Conference on VLSI Design (VLSID), pp. 535–540. IEEE (2015)
8. Winter, D.A.: *Biomechanics and Motor Control of Human Movement*. Wiley, Hoboken (2009)
9. Yang, X.-L., Hu, W., Yan, H., Xiao, M., Sun, X.: Research progress of noninvasive continuous blood pressure measurements based on pulse wave velocity. *Hangtian Yixue yu Yixue Gongcheng* **24**(6), 467–472 (2011)
10. Zhang, X.-Y., Zhang, Y.-T.: Model-based analysis of effects of systolic blood pressure on frequency characteristics of the second heart sound. In: 2006 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2006, pp. 2888–2891. IEEE (2006)



Some False ECG Waves Detections Revised by Fractal Dimensions

Ibticeme Sedjelmaci¹(✉) and Fethi Bereksi Reguig²(✉)

¹ Electrical Systems Engineering Department,
UMBB University, Boumerdès, Algeria
ibticeme_sedjelmaci@yahoo.fr

² Biomedical Engineering Laboratory, ABBT University, Tlemcen, Algeria
fethi.bereksi@mail.univ-tlemcen.dz

Abstract. In this paper, we used the fractal dimensions in ECG signals to identify the wave's detections failure. We check for the sensitivities and the importance of QRS and ST detection because different false wave's detections caused by the various types of interference and artefact are detected for some ECG signals presenting pathologies.

The fractal dimension is very sensitive to variations: if irregularities degree is great, the fractal dimension is high and vice versa. Different cases of pathologies decreased irregularities on the ECG signal so it causes a decrease in fractal dimension. However decreasing in irregularities is not necessarily pathological: a bad detection can also train it, because we have not the exact location of the beginning and the end of QRS complex or the end of the T wave, it causes a new variation in the dimension fractal which can skew the result.

For that reason and in order to get good results from algorithm detection, the fractal dimensions are calculated for each QRS complex and ST segment, for some ECG signals, to check their sensitivities in heart rate irregularities and false wave's detections so that make ECG interpretation system more effective.

Keywords: Electrocardiogram signals (ECG)
QRS and ST detection algorithm · Fractal dimension

1 Introduction

During recent years, there has been an increasing interest in non-linear analysis of physiological phenomena. This is due to the fact that many scientists are convinced of the non-linear nature of these physiological processes, which exhibit a complex behavior.

Fractal analysis [1] is a non-linear method developed to study the irregular complex objects [2, 3]. Because some natural phenomena showed a fractal behavior in biomedical domain, the fractal analysis has been applied with a certain success [4–6].

The Electrocardiogram (ECG) is a physiological signal generated by a complex system of self-regulation, its reading and interpretation [7] require a big habit that cannot be acquired by the doctor with a regular practice. Some electrocardiographs are delivered with software that can help to the diagnosis, but their approximate reliability cannot replace the doctor.

The Electrocardiogram (ECG) is a physiological signal generated by a complex system of self-regulation, its reading and interpretation [7] require a big habit that cannot be acquired by the doctor with a regular practice. Some electrocardiographs are delivered with software that can help to the diagnosis, but their approximate reliability cannot replace the doctor.

A normal ECG contains waves, intervals, segments, and one complex (Fig. 1). An interval in an electrocardiogram includes one segment and one or more waves and a segment is the region between two waves and they are important in diagnosis.

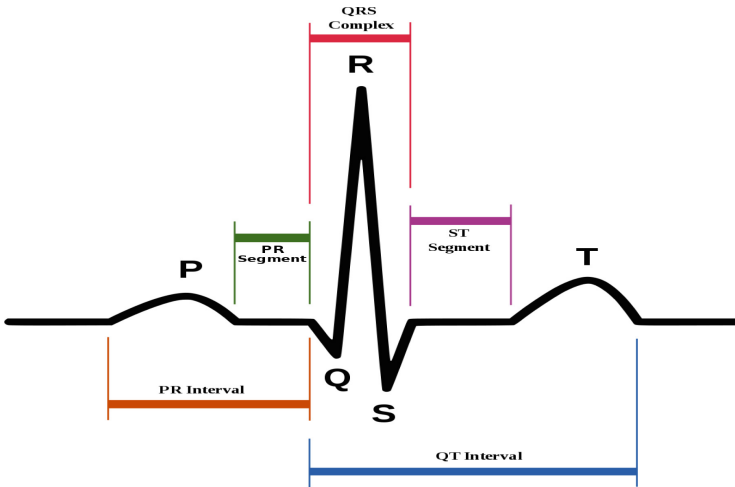


Fig. 1. Waves, intervals and the complex of ECG signal. (Color figure online)

Accurate QRS (or ST) detection is an important first step for the analysis of heart rate variability, there interpretation help to diagnosis the existence of pathologies. The QRS complex and ST segment can be detected by many methods and algorithms. The problem is: if there are some missed points in detection, we have incomplete or false information on the signal so we have to choose a sensitive method for checking. Fractal dimension is very sensitive to variations so we choose it to find some false detection.

2 Fractal Dimension

The well known part of fractals features is the fractal dimension. The various definitions of fractal dimensions are usually related the Hausdorff's original theoretical definition.

The Hausdorff dimension \dim_H is defined as the logarithmic ratio between the number N of an object's internal homotheties and the reciprocal of the common ratio r of this homothety:

$$\dim_H = \frac{\log(L_a)}{\log(1/a)} \quad (1)$$

Most of the time the problem in fractal analysis is the fractal dimension estimation. Various algorithms [8, 9] have been suggested for the practical estimation of the fractal dimension. These differences in algorithms appear because the Hausdorff dimension (Eq. (1)) is not computable in this form in most cases. Thus, the methods approximate it using different algorithms to estimate the parameter N .

We choose the regularization dimension [10] $\dim_R(\Gamma)$ which evaluate the regularity of a function f by measuring the dimension of its graph Γ . Let Γ be the graph of a bounded function f defined on some closed bounded interval $K \subset \mathbb{R}$. Let f_a be the projection of f on some approximation space at lower resolution $a > 0$ (such that $\lim_{a \rightarrow 0} f_a = f$), and define $L_a = \int_K \sqrt{1 + f_a'(t)^2} dt$, the finite length of the corresponding graph Γ_a . Then the regularization dimension is:

$$\dim_R(\Gamma) = 1 + \lim_{a \rightarrow 0} \frac{\log(L_a)}{-\log(a)}$$

f_a is the convolution of f with a Gaussian kernel.

3 Detection Algorithms

In this work, we are interested in assessing two well-known QRS and ST segments detection algorithms using fractal dimensions. These algorithms are respectively, the Pan and Tompkins algorithm [11] for QRS detection, and the “modified Chan and So” [12] algorithm for ST segment detection. The latter is based on the principle defined by Zhang et al. [13]. In both algorithms, the detection is carried out by first detecting the Q, the R peak, the S deflections and the end of the T wave, then generating the QRS and the ST segments. However in some cases these algorithms missed some of these points and therefore the detection is distorted [14]: the number of samples change and also the dimension.

4 Results and Discussion

The analyzed method determined the Regularization dimension (RD) for each QRS complex and ST segment detected in some signals (Tables 1 and 2). The results are showed in the drawn following graphs (Figs. 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13).

If the detections of some peaks are false, there location is very hard when we have 650000 samples, so we calculated the all dimensions and for these false detections designed by red squares in Figs. (1, 2, 3, 4, 5, 6, 8, 9, and 10) and by red circles in Figs. (12 and 13).

The dimensions are different. Thus we have exactly the information on our missed points.

The dimensions are calculated for all ST segments and QRS complexes in the signal, but we drawn just a part to have a good representation.

Table 1. Some annotations description of heartbeats.

Annotation	Description
N	Normal beats
APC	Atrial premature contraction
PVC	Premature ventricular contraction
RBBB	Right bundle branch block
Pb	Paced beats

4.1 QRS Complexes

The program represents the Q peaks are detected by red points and the S peaks by black points. The ECG signals are taken from MIT-BIH database and defined in Table 1:

Table 2. Some ECG signals and their features.

ECG signals	N	APC	PVC	Pb
“100”	☑	☑		
“101”	☑	☑		
“107”				☑
“200”	☑		☑	
“212”	☑		☑	
“209”	☑	☑	☑	

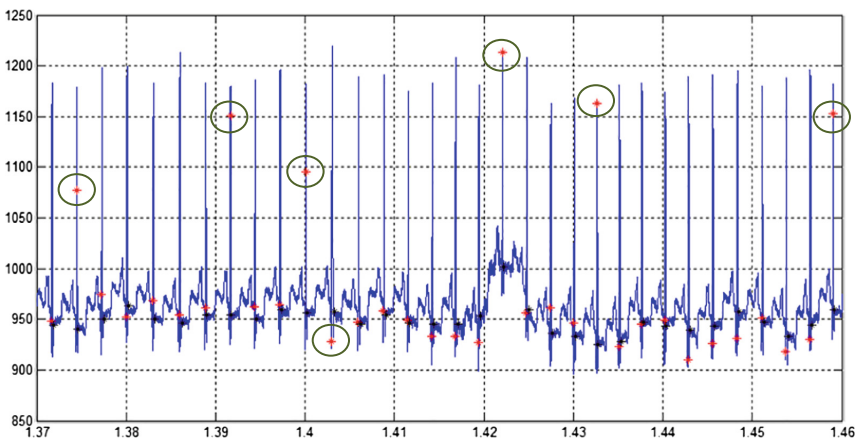


Fig. 2. Some QRS complexes detected for the signal number “100”. (Color figure online)

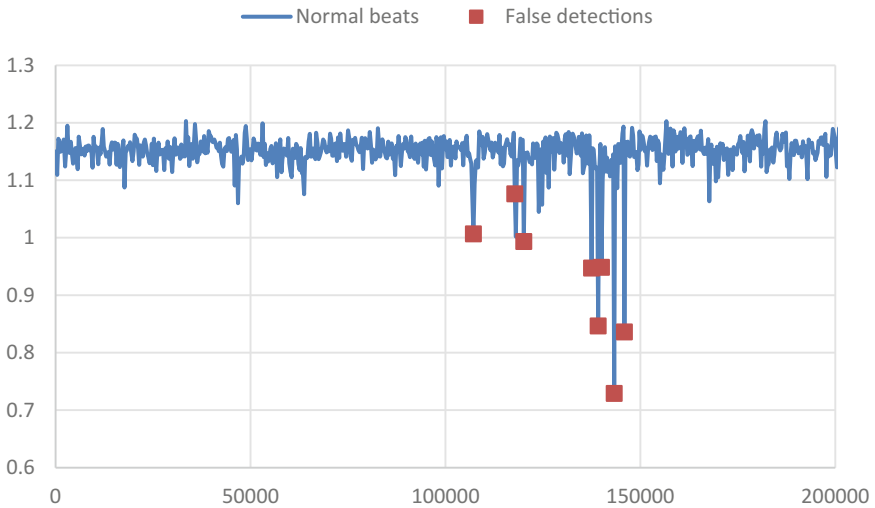


Fig. 3. Fractal dimensions for some QRS complexes of signal number "100". (Color figure online)

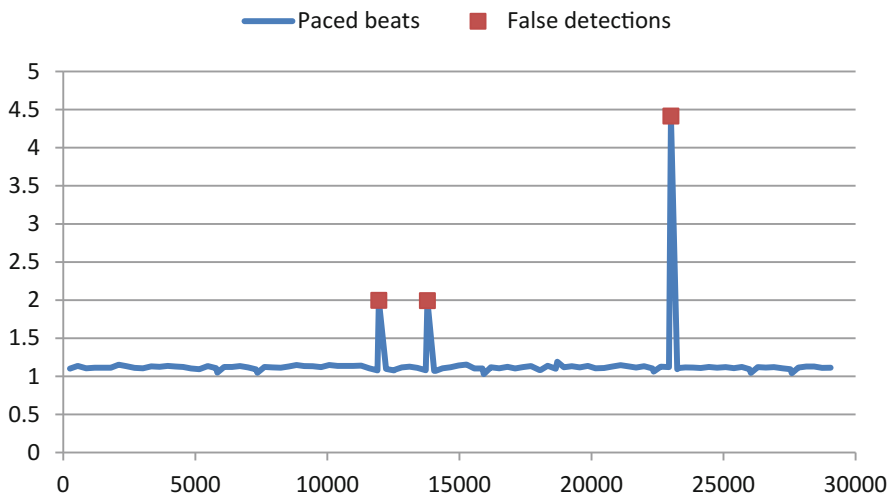


Fig. 4. Fractal dimensions for some QRS complexes of signal number "117". (Color figure online)

Physiological signals under healthy conditions may have a fractal temporal structure; the decrease of fractal behavior which causes the decrease of fractal dimension can be due to pathologies. A gradual decrease in the variability indicates a transition from a healthy state to a pathological state. However, in this case, it is not the reason of decreasing, since there are less singularities and less sample's number due to false detection.

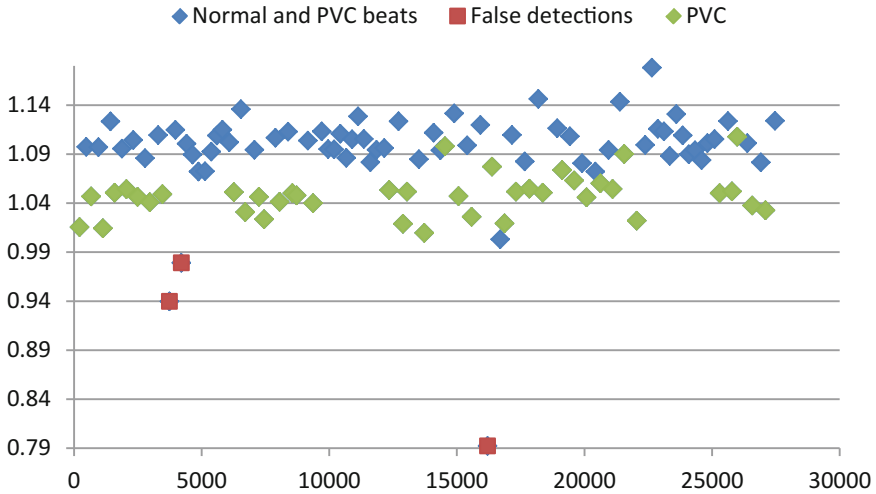


Fig. 5. Fractal dimensions for some QRS complexes of signal number “200”. (Color figure online)

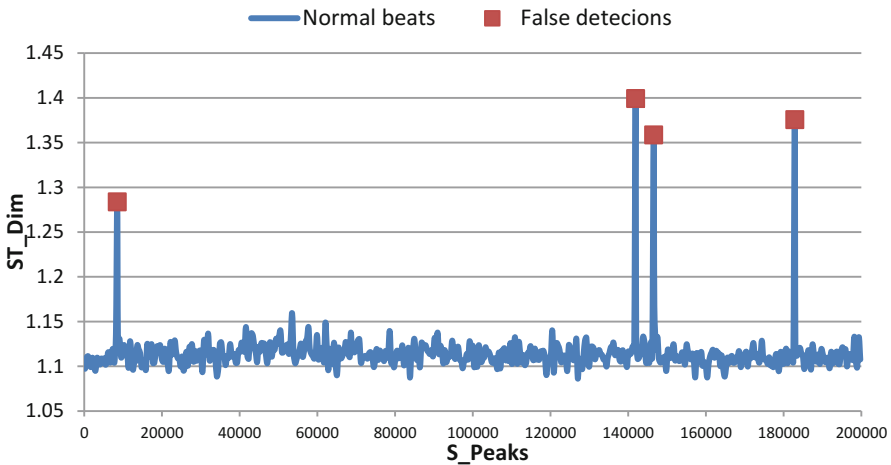


Fig. 6. Fractal dimensions for some ST segments of signal number “113”. (Color figure online)

4.2 ST Segments

In this case, the peaks S are in red color and the ends of T waves in black color (Fig. 4). We take those signals (Table 2):

ST_Dim_N-PVC: dimensions calculated for the normal segments just before the pathological ones (Table 3).

ST_Dim_PVC: dimensions calculated for the pathological segments.

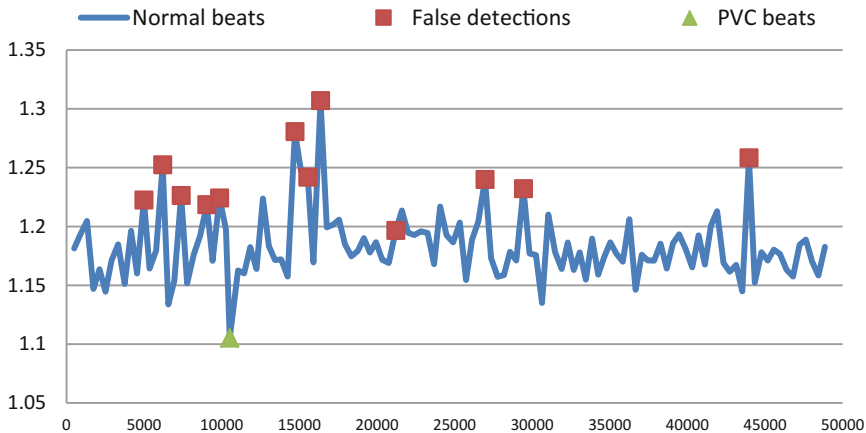


Fig. 7. Fractal dimensions for some ST segments of signal number “202”. (Color figure online)

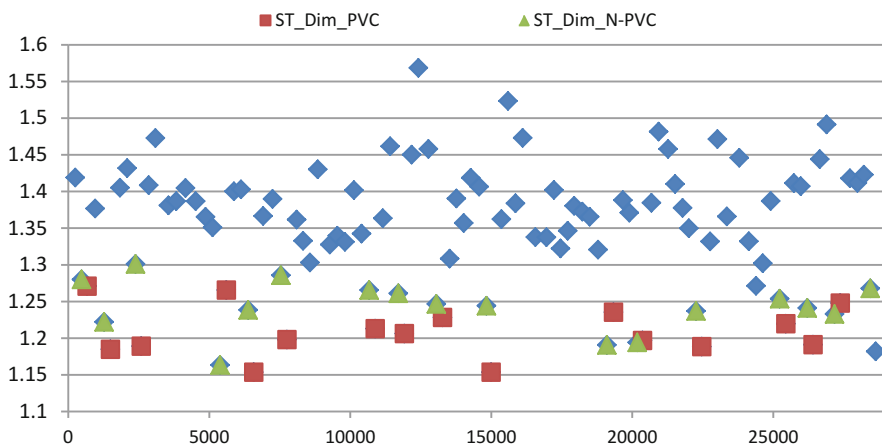


Fig. 8. Fractal dimensions for some ST segments of signal number “221”. (Color figure online)

In PVC, the QRS complex morphology changes and become larger than the normal one. On this signal (221), the morphology of the pathological QRS complex distorts the detection of the previous complex (N), the program missed the end of its T wave and the number of samples in this normal ST segment changes. So the fractal dimensions (ST_Dim_N-PVC) reduce compared with the others of normal ST segment but not under the values of pathological dimensions (ST_Dim_PVC). Therefore we will have information exactly on the existence of the pathology just before [16].

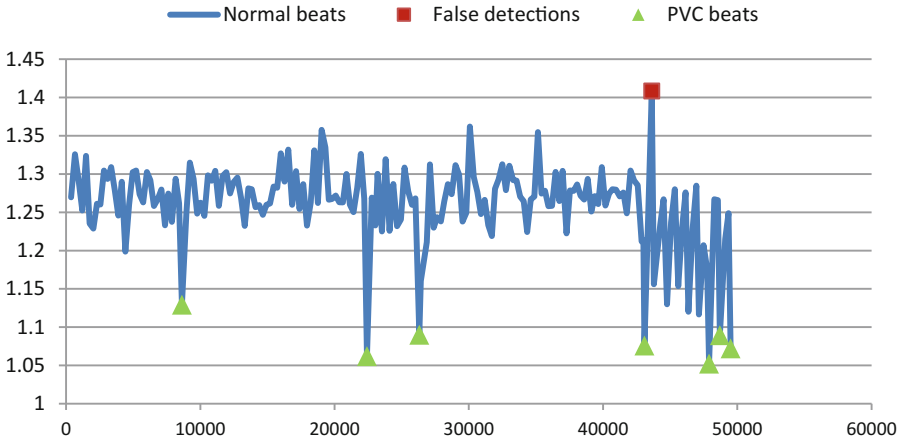


Fig. 9. Fractal dimensions for some ST segments of signal number “223”. (Color figure online)

In Fig. 10, the signal is fractioned in two distinguish groups: normal (N) and abnormal (RBB). In each group, the good detection of ST segments has a green circles in normal cases (N_ST) and a brown ones in RBB cases (RBB_ST).

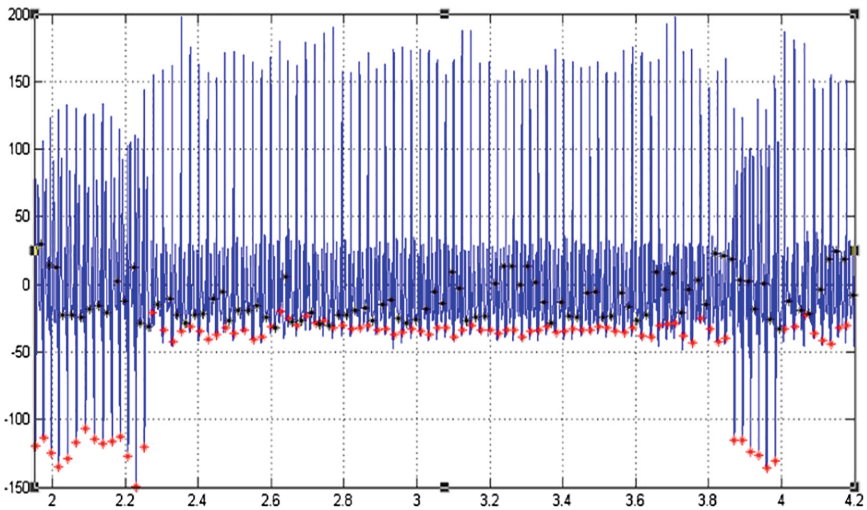


Fig. 10. Some ST segments detected for the signal number “212”. (Color figure online)

The fractal dimensions tell us separately about the pathological cases [15] of the signal (212), and the false detections on it (red squares).

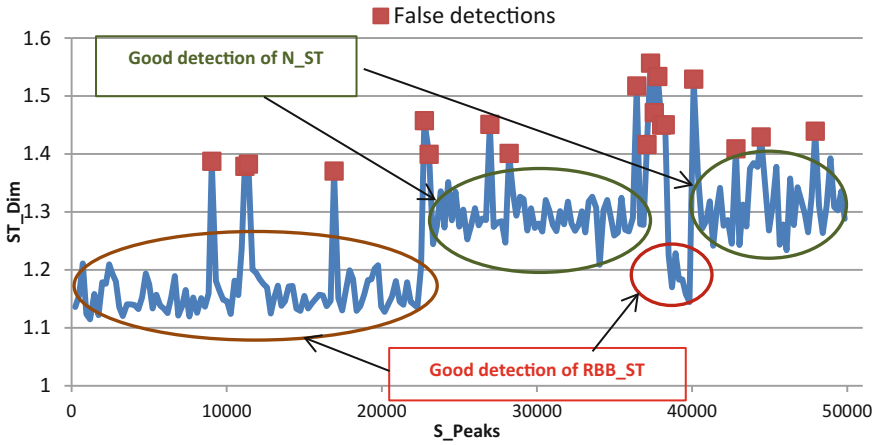


Fig. 11. Fractal dimensions for some ST segments of signal number “223”. (Color figure online)

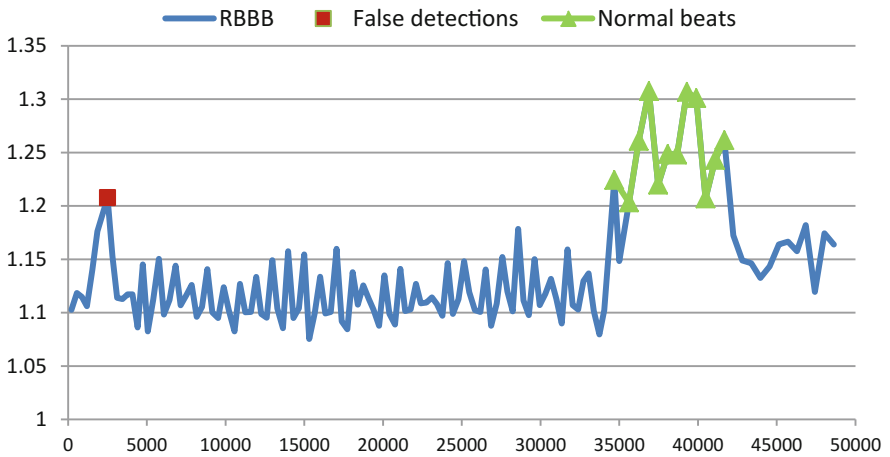


Fig. 12. Fractal dimensions for some ST segments of signal number “231”. (Color figure online)

The RBBB shows a difference on the T wave, so the detection of ST segments is very significant for pathology detection because the fractal dimension detect this difference on the waves.

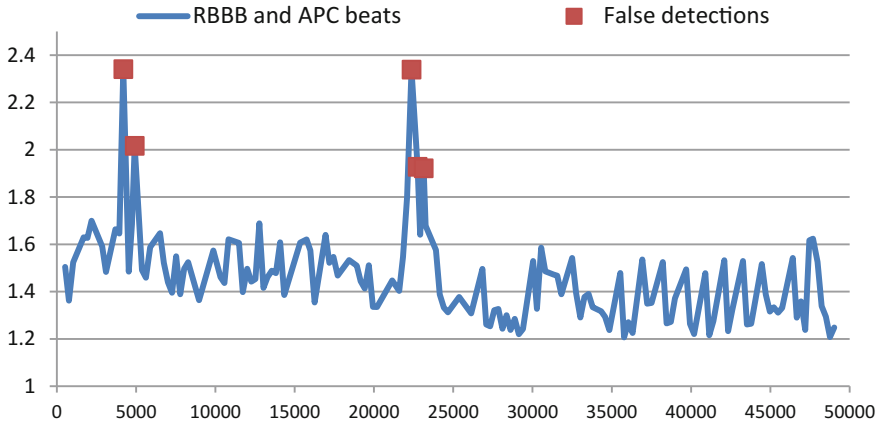


Fig. 13. Fractal dimensions for some ST segments of signal number “232”. (Color figure online)

Table 3. An others ECG signals and their features.

ECG signals	N	APC	PVC	RBBB	Pb
“113”	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
“202”	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
“212”	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	
“217”			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
“221”	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
“223”	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		

5 Conclusion

Fractal analysis studies signals with high variation in their regularities and fractal curves shows a very big degree of irregularity (singularity at each points), so fractal dimension is correlated to variations. The increase in dimension means high degree of irregularities.

In different cases of heart disease, we have a reduction in the irregularities of ECG signals and therefore it makes a reduction of fractal dimensions.

However all reductions of irregularities are not necessarily pathological: a false detection can also produce it, because in this case, the number of samples changes suddenly, reduces the irregularities and makes a change in the fractal dimension.

To confirm these results, we take more signals and have the same conclusion up to now.


References

1. Véhel, J.L.: Analyse Fractale: une nouvelle génération d'outils pour le Traitement du Signal. INRIA - Projet Fractales (2000)
2. Barrière, O., Véhel, J.L.: Local Hölder regularity-based modeling of RR intervals. Project APIS INRIA Saclay (2008)
3. Smrčka, P., Bittner, R., Vysoký, P., Hána, K.: Fractal and multifractal properties of heartbeat interval series in external states of the human organism. *Measur. Sci. Rev.* **3**, 13–15 (2003)
4. Lopes, R., Dubois, P., Bhourri, I., Akkari-Bettaieb, H., Mauouche, S., Betrouni, N.: La géométrie fractale pour l'analyse de signaux médicaux: état de l'art. *IRBM* **31**, 189–208 (2010)
5. Islam, N., Hamid, N.I.B., Mahmud, A., Rahman, S.M., Khan, A.H.: Detection of some major heart diseases using fractal analysis. *Int. J. Biom. Bioinfo. (IJBB)* **4**(2), 63 (2010)
6. Lee, J.S., Chang, K.S.: Applications of chaos and fractals in process systems engineering. *J. Process Control* **6**, 71–87 (1996)
7. Zhang, X.S., Zhu, Y.S., Zhang, X.J.: New approach to studies on ECG dynamics: extraction and analyses of QRS complex irregularity time series. *Med. Biol. Eng. Comput* **35**, 467–474 (1997)
8. Theiler, J.: Estimating fractal dimension. *J. Opt. Soc. Am. A* **7**(6), 1055–1073 (1990)
9. Jaffard, S.: Sur la dimension de boîte des graphes. *Math. Anal.* **326**, 555–560 (1998)
10. Roueff, F., Véhel, J.L.: A regularization approach to fractional dimension estimation. In: *Proceeding of Fractals 1998, Malta, October 1998*
11. Pan, J., Tompkins, J.: A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **32**, 230–236 (1985)
12. Chan, K., So, H.: Development of QRS method for real-time ambulatory cardiac monitor. In: *Proceedings of 19 Annual of International Conference IEEE EMBS, Chicago, USA*, pp. 289–292 (1997)
13. Zhang, Q., Manriquez, A.I., Médigue, C., Papelier, Y., Sorine, M.: An algorithm for robust and efficient location of T-wave ends in electrocardiogram. *Irisa Publication Interne1744* (2005). www.irisa.fr
14. Clifford, G.D.: ECG statistics, noise, artifacts, and missing data. In: *Advanced Methods & Tools for ECG Data Analysis*, chap. 3
15. Sedjelmaci, I., Bereksi Reguig, F.: La Théorie du Chaos et l'Analyse Fractale dans la Variabilité du Rythme Cardiaque. In: *Biomedical Engineering International Conference, BIOMEIC 2012, Tlemcen* (2012)
16. Sedjelmaci, I., Bereksi Reguig, F.: Detection of some heart diseases using fractal dimension and chaos theory. In: *The 8th International Workshop on Systems, Signal Processing and their Applications, WOSSPA 2013, Alger* (2013)

Challenges in Smart and Wearable Sensor Design for Mobile Health



Reconstruction of Equivalent Electrical Sources on Heart Surface

Galina V. Zhikhareva¹, Mikhail N. Kramm¹ , Oleg N. Bodin²,
Ralf Seepold³, Anton I. Chernikov¹, Yana A. Kupriyanova,
and Natalija A. Zhuravleva¹

¹ National Research University “MPEI”, Moscow, Russia
KrammMN@mail.ru

² Penza State University, Penza, Russia
iit@pnzgu.ru

³ Department of Computer Science, HTWG Konstanz, Konstanz, Germany
ralf.seepold@htwg-konstanz.de

Abstract. We consider the problem of increasing the informative value of electrocardiographic (ECG) surveys using data from multichannel electrocardiographic leads, that include both recorded electrocardiosignals and the coordinates of the electrodes placed on the surface of the human torso. In this area, we were interested in reconstruction of the surface distribution of the equivalent sources during the cardiac cycle at relatively low hardware cost. In our work, we propose to reconstruct the equivalent electrical sources by numerical methods, based on integral connection between the density of electrical sources and potential in a conductive medium. We consider maps of distributions of equivalent electric sources on the heart surface (HSSM), presenting source distributions in the form of a simple or double electrical layer. We indicate the dynamics of the heart electrical activity by the space-time mapping of equivalent electrical sources in HSSM.

Keywords: Multichannel · Electrocardiographic leads · Torso
Heart · Electric potential · Reconstruction · Equivalent electric sources
Maps of distributions

1 Introduction

The electrocardiographic method of heart investigation is one of the most common in cardiology. It is important to extract additional information on the electrical activity of the heart, using data from multichannel electrocardiographic leads, that include both recorded electrocardiosignals (ECS) and the coordinates of the electrodes placed on the surface of the human torso [1]. Since the potential distribution in a conducting medium is described by the Poisson equation, this leads to attenuation of the spatial harmonics in the potential map on the surface of the torso (BSPM), and therefore, leads to loses of small details in the image. It means, that the electric potential, recorded with some electrode, is an integral characteristic of bioelectric activity. It is difficult to link individual BSPM area with specific area of the heart epicardial surface. In this connection,

we are interested in solution of the inverse problem of electrocardiography, i.e. reconstruction of equivalent electric sources on heart surface, based on the records of multichannel electrocardiographic leads. As a result of such reconstruction, it is possible to analyze spatio-temporal maps of equivalent sources to reveal pathological regions of electrical activity of the heart. Reconstruction algorithms for distributions of simple and double layer sources density have been proposed and approved on simulated and real electrocardiosignals. These algorithms are based on algebraic method solution of the ECG inverse problem with Tikhonov regularization [2, 3]. Additional data for the algorithm are the coordinates of points on the epicardium surface. The determination of these coordinates requires considerable hardware and time costs and complicates the method [4, 5]. In this paper, we propose a more simple spectral method for determining the size and location of the epicardium. The spectral algorithm is based on the integral relation between the spatial spectra of the source and potential distribution.

We consider maps of distributions of equivalent electric sources on the heart surface (HSSM), presenting source distributions in the form of a simple or double electrical layer. We indicate the dynamics of the heart electrical activity by the space-time mapping of equivalent electrical sources in HSSM.

2 Algorithm of Reconstruction of Equivalent Electrical Sources

Relation between surface distributions of electrical sources and the measured electric potential in cases of simple and double sources layers follows from Poisson's equation:

$$\phi(\bar{r}) = \frac{1}{4\pi\sigma} \int_S \frac{\gamma_S(\bar{r}')}{|\bar{r} - \bar{r}'|} dS \text{ and } \phi(\bar{r}) = \frac{1}{4\pi\sigma} \int_S D_S(\bar{r}') \frac{\partial}{\partial n} \left(\frac{1}{|\bar{r} - \bar{r}'|} \right) dS,$$

where γ_S - density of unipolar moment, D_S - density of dipolar moment, $|\bar{r} - \bar{r}'|$ - distance from a measurement point (radius-vector \bar{r}) to the current point of integration (radius-vector \bar{r}') on quasi-epicardium surface S , $\partial/\partial n$ - a derivative in the direction of a normal to this surface.

The task is reduced to the decision of systems of the linear algebraic equations

$$\phi_k = \sum_l A_{\gamma kl} \gamma_{Sl} \text{ and } \phi_k = \sum_l A_{Dkl} D_{Sl} \quad (1)$$

with matrixes of coefficients

$$A_{\gamma kl} = \frac{1}{4\pi\sigma} \frac{\Delta S_l}{|\bar{r}_k - \bar{r}'_l|} \text{ and } A_{Dkl} = \frac{1}{4\pi\sigma} \frac{\Delta S_l \cos \alpha_{lk}}{|\bar{r}_k - \bar{r}'_l|^2}$$

where ΔS_l - the square of a discrete quasi-epicardium element with number l , α_{lk} - angle between a vector of the dipolar moment D_{Sl} and a straight line connecting elements with numbers l and k on surface S and a torso.

The decision of systems of the Eq. (1) belongs to incorrect tasks, therefore the method of A.N. Tikhonov regularization was applied to find the normal decision: we need to search of a vector z^α , minimizing functional

$$M^\alpha[z, \phi] = \|Az - \phi\|^2 + \alpha_a \|z\|^2, \quad \alpha_a > 0, \quad (2)$$

with regularization parameter α_a , defined on discrepancy, i.e. from a condition $\|Az^\alpha - f\| = \delta_1$, where δ_1 – total mean square deviation of the measured potentials from ones, calculated by means of the found decision $\gamma_S = z^\alpha$. Minimization of functional (2) is reduced to the decision of system of the linear algebraic equations

$$\alpha_a z_k^\alpha + \sum_j C_{kj} z_j^\alpha = b_k,$$

with

$$C_{kj} = \sum_i A_{ik} A_{ij}, \quad b_k = \sum_i A_{ik} \phi_i.$$

3 Interpolation of Torso Surface and the Surface Potential Map

Electric potentials were registered by means of the multielectrode measuring system shown in Fig. 1 [3]. Electrodes were installed on elastic belts placed on the torso. This system is suitable for both men and women with small breasts. For women with large breasts, the position of the electrodes on the front of the torso can be adjusted. Coordinates of electrodes were defined by the separate program, processing digital photos of a torso with electrodes. Basic data for reconstruction are values of potentials in rather large number of points of a surface of a torso and coordinate of these points (about 103 points). Therefore, it is necessary to carry out interpolation procedures for determination of values of coordinates and potentials in intermediate points of a torso.

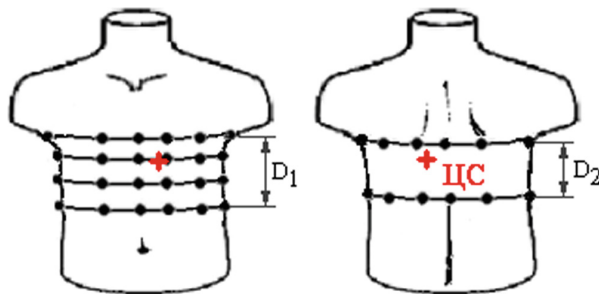


Fig. 1. System of electrode leads. ИЦ – heart center

At the first step of interpolation we determined coordinates of intermediate points of torso surface using coordinates of electrodes, i.e. carried out reconstruction of torso surface. At the second step we found the interpolated values of potentials in the found intermediate points. In Fig. 2 we presented surface potential maps (BSPM) for the time points, corresponding to P-, R- and T-peaks of a cardiocycle (Fig. 3). Figures 2 and 3 are obtained for real electrocardiosignals. On Fig. 2, the magnitude of the electrical potential is displayed using the pseudocolor scale: the largest positive potential value is displayed by the largest saturation of red color, and for the negative potential – by the saturation of blue color. To reduce the noises level up to 1-2 μV , we used synchronously accumulated recordings of the cardiocycles in combination with digital filtration.

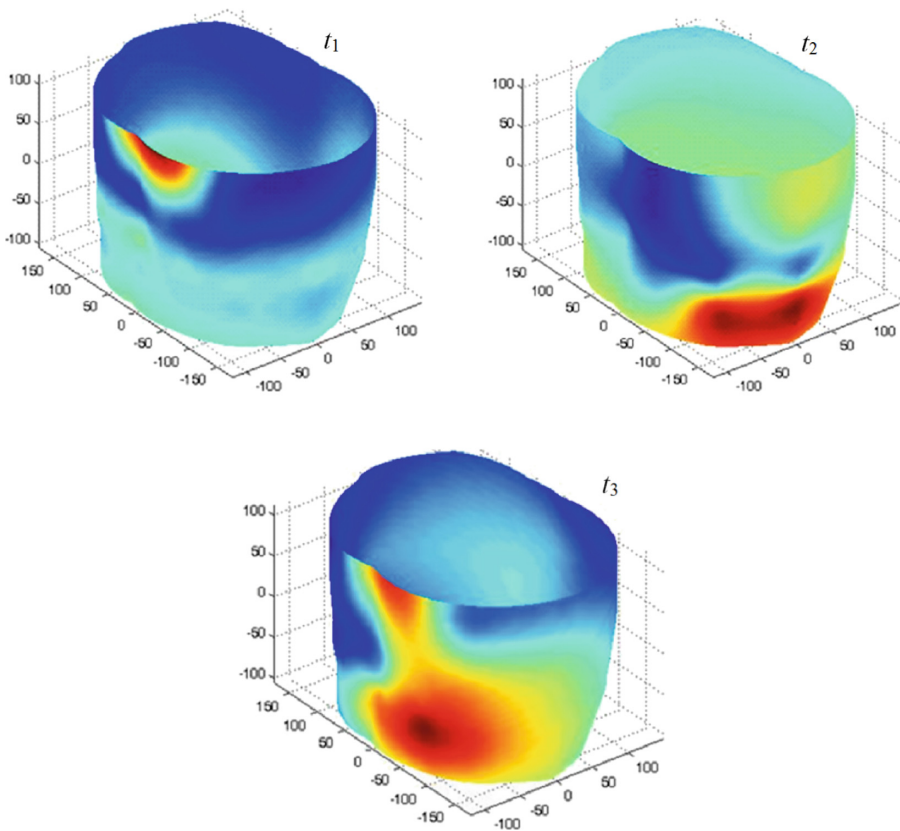


Fig. 2. Surface potential maps for time points t_1 , t_2 and t_3 (Fig. 3); spatial coordinates are measured in millimeters (Color figure online)

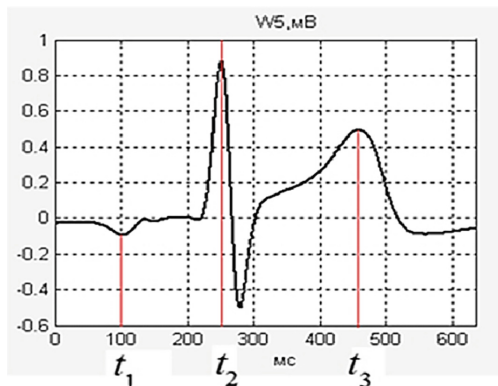


Fig. 3. Time points of cardiocycle, the magnitude of the electrocardio signal is measured in millivolts

4 Determination of Position and Size of Quasi-epicardium

It is suitable to reconstruct the equivalent electrical sources, defined on a closed surface surrounding the myocardium (or quasi-epicardium sphere map [1]).

For definition of position and size of quasi-epicardium we carried out reconstruction of area of electrical sources in the frontal and sagittal planes. We used the spectral algorithm of reconstruction, which follows from relation between spatial spectrum of distributions of electrical sources density and spectrum of potentials in some planes divided by distance of z_0 [6, 7]:

$$\gamma_S(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma_S(k_x, k_y) W_\alpha(k_x, k_y) e^{i2\pi(k_x x + k_y y)} dk_x dk_y$$

where $\Gamma_S(k_x, k_y) = F(k_x, k_y)\Phi(k_x, k_y)$ – spatial spectrum of distributions of electrical sources $\gamma_S(x, y)$, $\Phi(k_x, k_y)$ – spatial spectrum of potentials;

$$F(k_x, k_y) = 4\pi\sigma\sqrt{k_x^2 + k_y^2}e^{2\pi z_0\sqrt{k_x^2 + k_y^2}}$$

- the frequency transfer coefficient, deriving from properties of Laplace operator;

$$W_\alpha(k_x, k_y) = \begin{cases} 0.5 + 0.5 \cos\left(\frac{\pi\sqrt{k_x^2 + k_y^2}}{k_\alpha}\right), & \sqrt{k_x^2 + k_y^2} \leq k_\alpha \\ 0, & \sqrt{k_x^2 + k_y^2} > k_\alpha \end{cases}$$

- regularizing function of frequency window with band $k_\alpha = (\Delta\Phi + \Delta\Gamma_S)/2$, $\Delta\Phi$ и $\Delta\Gamma_S$

- width of distribution of the potential and the source spectra. For application of spectral algorithm we recalculated surface potential maps (SPM's) to frontal plane and sagittal plane, presented in Fig. 4. Here it is shown, in what planes distributions of sources were reconstructed.

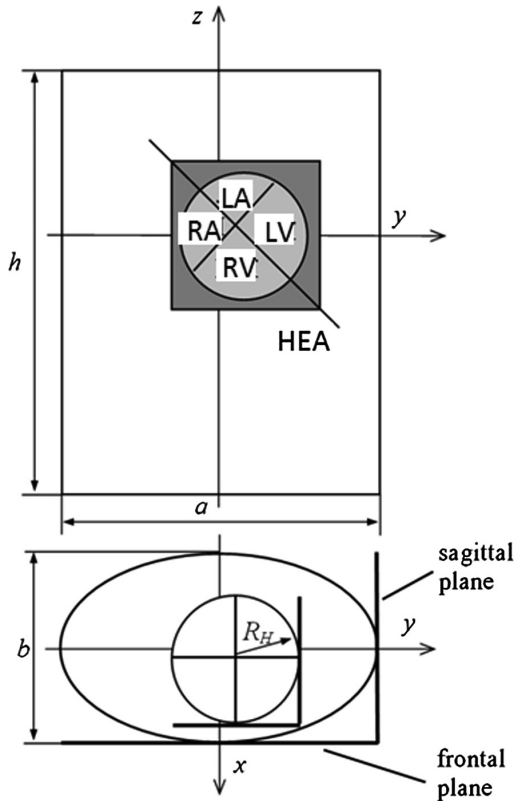


Fig. 4. Planes for potentials and sources. RA and LA – the right and left atriums, RV and LV – the right and left ventricles, HEA – electrical axis of the heart

As an example, distribution of a source in the frontal plane is presented in Fig. 5.

For determination of position and size of quasi-epicardium it is necessary to summarize the reconstructed distributions of sources γ_S for all cardiocycle peaks on time points (separately on frontal and on sagittal planes) [3]. Then, based on summarized distribution γ_S in these planes, matrixes of spatial location of a source were formed:

$$I(i,j) = \begin{cases} 1, & \gamma_{\Sigma}(i,j) \geq 0.5\gamma_{\Sigma\max}, \\ 0, & \gamma_{\Sigma}(i,j) < 0.5\gamma_{\Sigma\max}. \end{cases}$$

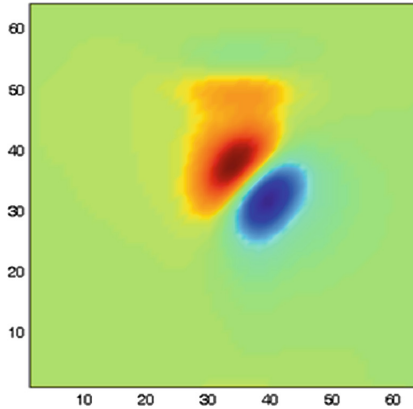


Fig. 5. Example of distribution of source density γ_S in the frontal plane

Having matrixes of spatial location (Fig. 6) and knowing numbers of extreme nonzero elements, we determined average borders of area of electrical sources, and from here we found position and the size of quasi-epicardium.

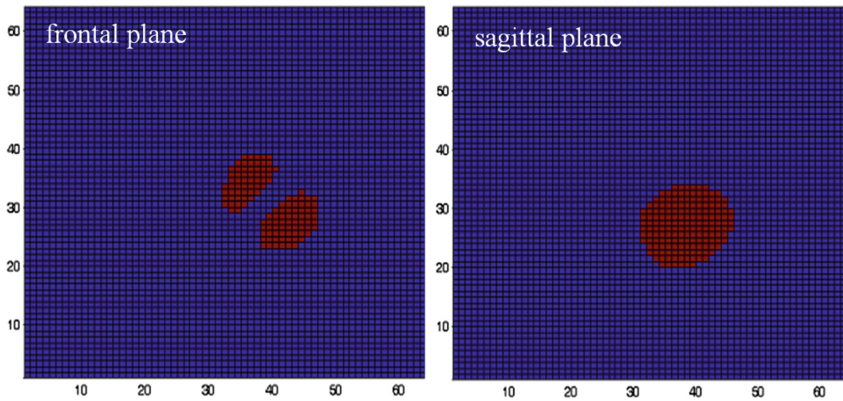


Fig. 6. Matrixes of spatial location of a source in frontal and sagittal planes

5 Results of Reconstruction

The presented methods were used for reconstruction of equivalent electrical sources on spherical quasi-epicardium, based on real multichannel records of electrocardiosignals. The received surface source distributions (HSSM) were represented on plane projection of the sphere quasi-epicardium (Fig. 7). In Fig. 8 we represent HSSM's for the time points, corresponding to tops of P, R and the T-peaks of a cardiocycle (Fig. 3). For convenience of visualization the scale of pseudo-colors on maps of Fig. 8 fits to different range of sources density.

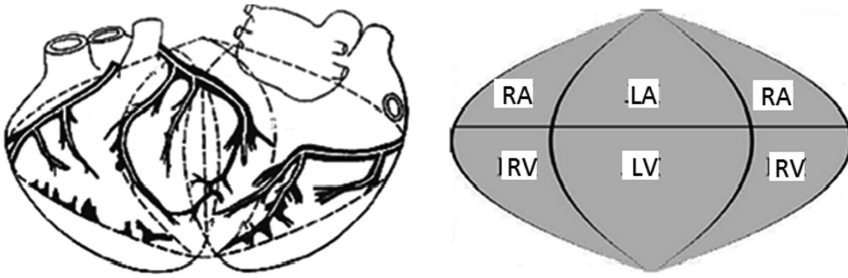


Fig. 7. The plane projection of quasi-epicardium surface; areas of auricles and ventricles. RA and LA – the right and left atriums, RV and LV – the right and left ventricles

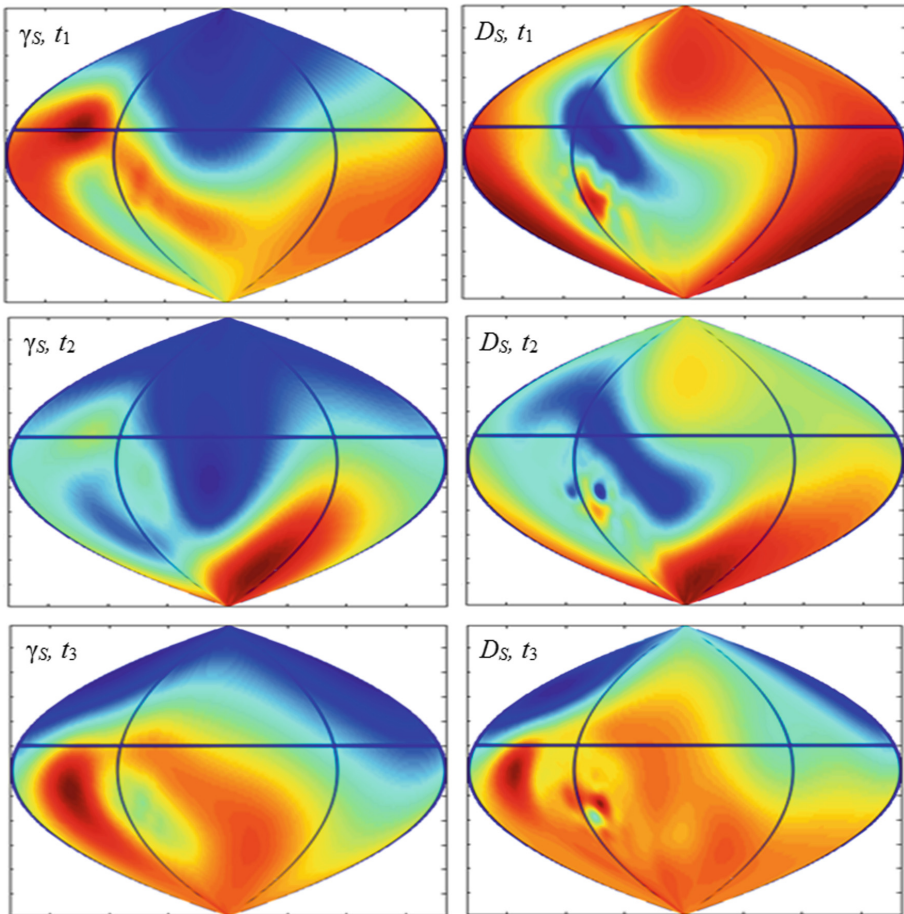


Fig. 8. Surface source distributions for the time points, corresponding to tops of P, R and the T-peaks of a cardiocycle (Fig. 3)

In general, the source distributions for the cases of a simple and a double layer correspond to each other. The density of a simple layer reflects the density of effluent currents, and the density of a double layer reflects the behavior of the transmembrane potential in depolarization and repolarization of the heart. We observe the smaller details in the distribution of the sources of the double layer.

In Fig. 9 we represent simple layer HSSM's for various time points of a QRS complex. For the distributions of the right column in Fig. 9, a scale with uniform pseudocolor saturation is chosen, when the range of source density values for the R-peak moment is described by the full saturation range. For the distributions of the left column in Fig. 9, a scale with an uneven pseudocolor saturation is chosen when a narrow range of source density values is described by the full saturation range. This increases the convenience of observing the boundaries of regions with various electrical activity of the heart. For example, for time points 3 and 4, corresponding to the region of the top of the R-wave, we observe the greatest electrical activity (depolarization) of the left ventricle and a small region in the upper part of the right ventricle.

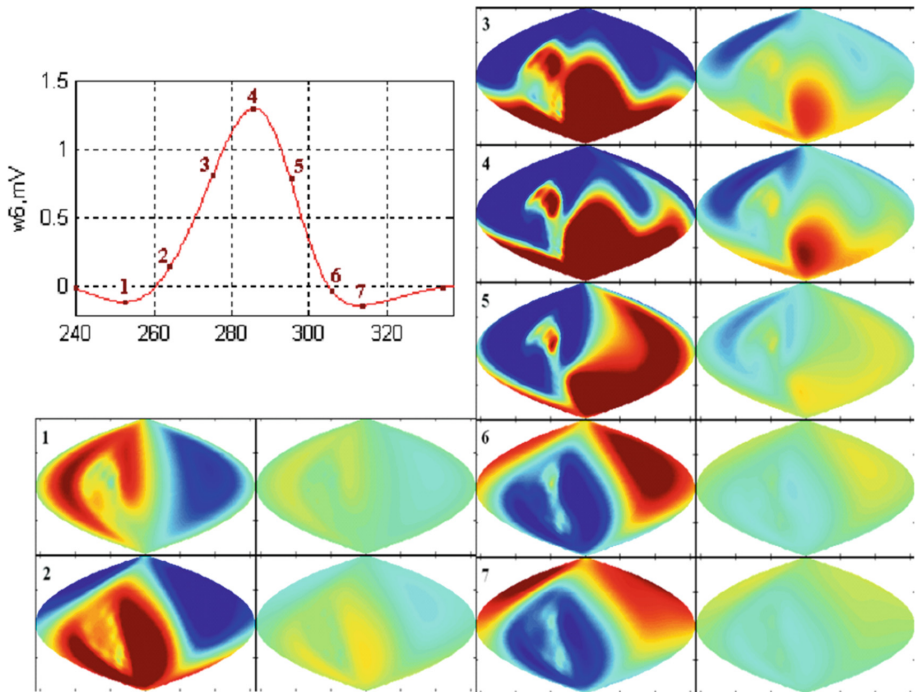


Fig. 9. Surface source distributions for the time points, corresponding to various time points of a QRS complex

6 Conclusion

The presented algorithms allow to indicate the dynamics of the heart electrical activity by the space-time mapping of equivalent electrical sources and can be recommended for diagnostics of conductivity processes in electrocardiology.

References

1. Titomir, L.I., Trunov, V.G., Aidu, E.A.I.: *Noninvasive Electrocardiotopography*. Nauka, Moscow (2003)
2. Macfarlane, P.W., van Oosterom, A., Pahlm, O., Kligfield, P., Janse, M., Camm, J.: *Comprehensive Electrocardiology*, 2nd edn, p. 2291. Springer, London (2011). (Chapter 9)
3. Zhikhareva, G., Kramm, M.: *Reconstruction of Current Sources of Heart in the ECG Inverse Problem*. LAP LAMBERT Academic Publishing GmbH & Co. KG, Saarbrücken (2012)
4. Revishvili, A.S., Kalinin, V.V., Lyadzhina, O.S., Fetisova, E.A.: Verification of a new technique for non-invasive electrophysiological examination of the heart, based on solving the inverse problem of electrocardiography. *Her. Arrhythm.* **51**, 1–13 (2008)
5. Potyagaylo, D., Cortes, E.G., Schulze, W.H.W., Dössel, O.: Binary optimization for source localization in the inverse problem of ECG. *Med. Biol. Eng. Comput.* **52**, 717–728 (2014)
6. Zhikhareva, G.V., Kramm, M.N.: Feasibility study the localization of pathological areas of myocardium. *Mechatron. Autom. Control* **2**, 46–51 (2007)
7. Filonov, D.V., Vinokurov, D.S., Kramm, M.N., Zhikhareva, G.V.: Reconstruction of myocardial current sources from measured surface potentials. *Meas. Tech.* **52**, 1015–1019 (2009)



WearIT - A Rapid Prototyping Platform for Wearables

Isabel Leber^(✉) and Natividad Martínez Madrid^(✉)

Reutlingen University, Alteburgstraße 150, 72762 Reutlingen, Germany
isabel.leber@student.reutlingen-university.de,
natividad.martinez@reutlingen-university.de
<http://iotlab.reutlingen-university.de>

Abstract. Rapid Prototyping Platforms reduce development time by allowing quick prototyping of a prototype idea and achieve more time for actual application development with user interfaces. This approach has long been followed in technical platforms, such as the Arduino. To transfer this form of prototyping to wearables, WearIT is presented in this paper. WearIT consists of four components as a wearable prototyping platform: (1) a vest, (2) sensor and actuator shields, (3) its own library and (4) a motherboard consisting of Arduino, Raspberry Pi, a board and a GPS module. As a result, a wearable prototype can be quickly developed by attaching sensor and actuator shields to the WearIT vest. These sensor and actuator shields can then be programmed through the WearIT library. Via Virtual Network Computing (VNC) with a remote computer, the screen contents of the Raspberry Pi can be accessed and the Arduino be programmed.

Keywords: Wearables · Rapid Prototyping Platform · Shields
Sensors · Actuators · eHealth · mHealth

1 Introduction

In software development, prototypes are used to quickly implement a first version of the application and get early feedback on this application from, for example, users or customers. This prototyping enables the developer to obtain feedback on the suitability of his idea at an early stage of development. A big trend in computer science is the use of wearables. Wearables are smart watches, wristbands, headsets, jewelry, textiles, and smart garments, among others [15]. These are mostly used to improve communication, monitor body functions, monitor the wearer's environment, and alert the wearer in the form of beeps or haptic signals [15].

The development of such wearables can be very time consuming and costly. Therefore, to test a development idea very early, a wearable prototyping platform is needed. With this, the basic ideas of a later application can be tested very quickly and easily. In the technical field there are already platforms, such as Arduino, that support the developer in the rapid development of a first prototype. The microcontroller can be supplemented by further functionalities in

the form of shields. Shields are printed circuit boards that are plugged onto the microcontroller and thus expand its functionality. This principle will be transferred to WearIT as a wearable rapid prototyping platform.

1.1 Goal

A number of technical platforms for rapid prototyping are already being used in certain application domains. However, wearables developers so far face the problem that rapid prototypes are difficult to create. This is because the electronic components must be attached to the body. For this implementation, electronics usually have to be sewn up in a very time-consuming manner and connected to a microcontroller or computer.

The goal of this work is therefore the development of a rapid prototyping platform for wearables. With the help of this platform, later application functions should be tested easily and quickly on a vest.

1.2 Layout

The rest of this paper is divided into a total of four parts. In the first part, the basic requirements are formulated based on WearIT's objectives as a rapid prototyping platform. In the following part discussing the state of the art, existing prototyping platforms will be presented and the distinction of this approach to the related work highlighted. The third part of this work describes the technical implementation of WearIT. First, the structure of the components is modeled and the structure of the wearable shields is defined. Then the structure of the mainboard and the WearIT library will be described. Finally, a conclusion and an outlook on possible future work will be given.

2 Requirements for WearIT

WearIT is a rapid prototyping platform designed to test the feasibility of new ideas for wearables. The developers should be able to quickly test new ideas based on a prototype. To accomplish this, WearIT must meet the following requirements:

- Sensors or actuators should be able to be quickly mounted on the platform.
- It should also be possible to extend the platform with future sensors and actuators.
- The shields with the sensors and actuators should be able to be placed freely.
- The wearable rapid prototyping platform should be able to be adapted to the clothing size of the respective wearer.
- A library should be able to allow quick and easy access the existing sensors and actuators.

WearIT should not be limited to an application domain, but rather be used for all interactive projects. Particularly, because of the possibility to develop prototypes with direct body contact to measure body signals and user movement, WearIT is specially suitable for eHealth and mHealth applications.

3 State of the Art

In the following, existing prototyping platforms will be discussed, the differences to this work will be outlined and related work will be presented.

3.1 Prototyping Platforms

In technical development, ideas and solutions are tested using prototypes [6]. Moreover, prototyping is used to communicate ideas and solutions to, for example, future users [6]. Typically, the focus is first set on the functionality and the form of the prototype is initially neglected [8]. Prototypes do not only test solutions, they are also used to generate new ideas [6]. In particular, the goal in rapid prototyping is to be able to quickly test ideas and solutions through a prototype. According to Hodge et al. the relevance of technical prototyping platforms will become more important in the future [8]. There are already some available platforms for the generation of technical prototypes.

- Arduino: One of the most popular prototyping platforms is the Arduino [3]. The Arduino is an I O board with a microcontroller [8] and a large selection of stackable shields [12].
- PC-104 and Beaglebone: The PC-104 and Beaglebone are other platforms that can be expanded with stackable shields and many wireless sensors and actuators [8].

Shields can be connected by stacking on these platforms [8] and are therefore very restricted in their use.

A rapid prototyping platform that is not limited to the stacking of shields is the Calder Toolkit. The Calder Toolkit is a platform to create fast, working prototypes for physical user input [11]. To accomplish this, the Calder Toolkit provides reusable small input and output components.

There is an interface available for these components to the Calder Toolkit development environment. An external PC is used for the calculations. The components used are all equipped with a microcontroller. Communication is via cable or wireless. For wired communication, a Universal Serial Bus (USB) is used. For wireless communication, an uplink transceiver is used.

Another rapid prototyping platform is OpenCapSense. OpenCapSense is an open source toolkit that allows developers to quickly prototype user interfaces [7]. Capacitive sensors are used for the user interfaces. OpenCapSense consists of a microcontroller board and eight sensor channels that connect sensors to a USB cable. In addition, OpenCapSense includes the software framework SenseKit, with which the input data can be analyzed.

Another modular rapid prototyping platform is Microsoft .NET Gadgeteer [8]. This platform contains a central motherboard with microcontroller. The prototypes are developed using modules that are connected to the motherboard via 10-pin ribbon cables. Gadgeteer is characterized by the fact that different shields using the ribbon cable can be easily connected to each other.

MySignals is another prototyping platform with the application domain of vital biosignals [1]. MySignals offers the possibility to rapidly develop medical and eHealth applications. For this purpose, the 16 sensors from MySignals can be used, or even own sensors can be connected. The 16 MySignals sensors are exclusively sensors for the measurement of biosignals. MySensors includes an Arduino compatible motherboard with an Atmega 328 microprocessor. The sensor data are connected to the mainboard via a jack plug (up to five pins). In addition, it is possible to connect sensors via Wifi, BEE or BT2.0. The data can be viewed via an integrated TFT screen or on mobile devices such as mobile phones or notebooks.

The LilyPad Arduino is the last prototyping platform presented here, which specializes in the application domain of wearables. The LilyPad is an Arduino that allows a technical development based on textiles [5]. The LilyPad Arduino consists of a motherboard with a microcontroller to which various LilyPad shields with sensors and actuators can be connected [4]. These shields, as well as the motherboard, have eyelets that allow them to be sewn to textiles. The connection between the shields and the motherboard is realized via conductive threads. Through this sewn connection there is both an electrical and a physical connection. Through this connection, the power supply and the data exchange takes place.

3.2 Comparison of Existing Platforms and WearIT

WearIT differentiates itself as a wearable rapid prototyping platform from other platforms, such as Arduino or PC-104, by the fact that WearIT is a specific platform as a wearable. Therefore, stacking of shields as used in Arduino or PC-104 is unsuitable for WearIT.

Likewise is communication via a USB connection, as used in the Calder Toolkit or OpenCapSense, unsuitable for WearIT. Due to the USB connection, a simple and fast extension of WearIT with new sensors and actuators would not be possible. This would require each sensor and actuator to be equipped with a USB functionality that they do not have. In addition, these platforms are not prototyping platforms for wearables. Even the prototyping platform Microsoft .Net Gadgeteer is not a prototyping platform for wearables. One thing in common with WearIT is its expandability with new sensors and actuators.

MySignals is a prototyping platform for medical applications. Since WearIT is not restricted to medical applications in the application domain, it already differs from MySignals. In addition, MySignals uses only sensors for the measurement of biosignals, while WearIT also includes the use of actuators. The communication of the sensors of MySignals is realized via jack plug, therefore MySignals is limited to max. 5 pins in the communication. In order to ensure flexibility for the application of new sensors and actuators, WearIT is not limited to a 5-pin connection.

Like WearIT, the LilyPad Arduino is being used as a prototyping platform for wearables. However, WearIT differs from LilyPad in that WearIT is a rapid

prototyping platform. Fast prototyping is not possible with LilyPad because of the need to sew in the new circuits.

3.3 Related Work

Sensors and actuators on clothing are already being used for a variety of applications. In the following, two application areas will be exemplified by scientific work.

An area is the medical domain. The user's biosignals are monitored while he is at home as with Hung et al. [9]. Here, the blood pressure and the heartbeat is detected by a finger ring sensor.

Another work with a medical background is the Georgia Tech Wearable Motherboard [13]. This is a sensor vest that can be used in military context. The health status of soldiers can be monitored.

Another example of a piece of clothing for the monitoring of biosignals is the Vivo Metrics Life Shirt [17]. With the help of this garment plethysmographic, acceleration and ECG data can be captured.

The second application area shown here is the security area. Here, Jutila et al. [10] provide a child safety vest. This vest contains a GPS sensor, temperature sensors and an acceleration sensor. This is to record information about the position and well-being of the child.

The safety vest approach is also pursued by Aleksy et al. [2]. They provide a safety vest in the context of human-machine applications. It uses sensors to monitor the environment and the wearer's vital signs.

4 Technical Implementation

The technical implementation of WearIT is divided into three major parts:

During the design of the structure of the components, the vest was selected as a platform and the overall structure of the application was specified. The sensor and actuator shields and their connection to computer and microcontroller are shown in the overall structure.

For the implementation of wearable shields, basic structures were first selected and defined. In addition, the component communication was defined and the motherboard was developed.

To access WearIT's sensor and actuator shields, a library of the most important control features of wearable shields was created.

4.1 Structure of the Components

A garment is required as basic component of the WearIT prototyping platform for quick and flexible mounting of sensor and actuator shields. For this purpose, specific requirements are defined.

To meet wearability [14] requirements, the garment should be comfortable, easy to wear and take off, allow enough freedom of movement, and provide good

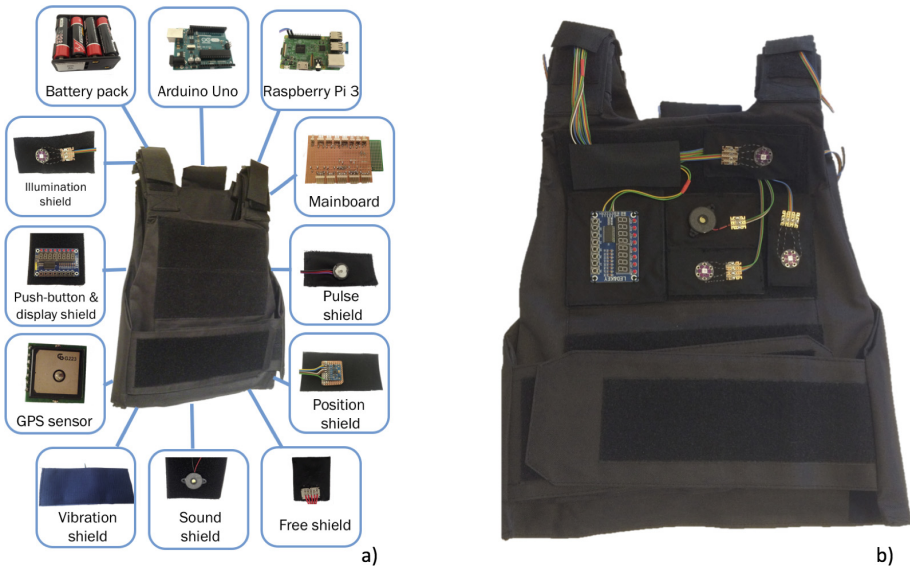


Fig. 1. Components of the application (a) and final application (b)

access to the body [14]. Additional requirements include being able to change the size of clothing to fit the wearer and the garment's stability to attach electronics to it.

For the fulfillment of these requirements, a vest has been chosen which can be flexibly adjusted in length in the shoulder area and can be adjusted in width by means of a Velcro fastener (see Fig. 1). For the expansion of sensors and actuators, sensor and actuator shields were created. As sensor shields, WearIT has a pulse shield for measuring the pulse, a position shield for measuring translation and rotation, a push-button and display shield with programmable buttons and a GPS sensor (see Fig. 1a). As actuator shields, WearIT has a sound shield for sound delivery, illumination shields, and vibration shields (see Fig. 1). For the flexible extension of WearIT, free shields are provided for sensors and actuators (see Fig. 1a).

In addition, WearIT has an Arduino Uno as microcontroller and a Raspberry Pi 3 model B as a computer (see Fig. 1a). The power source is a 9V battery pack (see Fig. 1a). Depending of the usage, a running time up to two hours was observed in practice.

In (see Fig. 1b) you can see the final application of WearIT. The sensor and actuator shields are attached with Velcro fastener on the vest and connected via cable connection to the mainboard.

In Fig. 2 WearIT was modeled as a complete system with all components using FMC (Fundamental Modeling Concepts¹). There, the interactions of the

¹ <http://www.fmc-modeling.org/>.

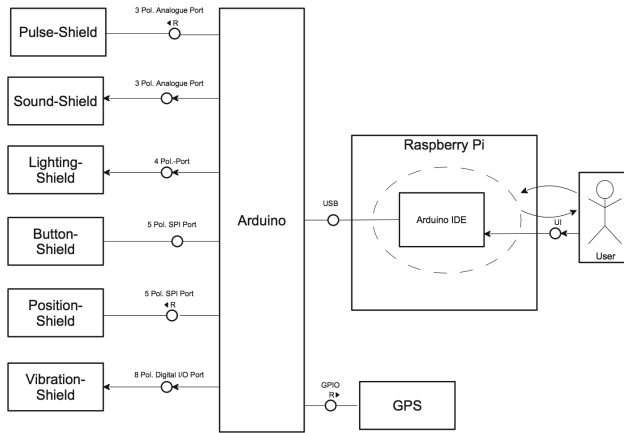


Fig. 2. Structure of the overall system

different components are displayed one below the other. The sensor and actuator shields, Arduino and Raspberry Pi as well as the Arduino IDE are represented as components (see Fig. 2).

The Raspberry Pi includes the Arduino IDE as an active system component. Since the Arduino IDE can change over the term of WearIT, this is a structural variance. The Arduino IDE is captured as a subsystem and memory. This memory can be modified by the users programming.

Communication between the Arduino IDE and the Arduino takes place via a bidirectional communication channel in the form of a USB connection (see Fig. 2). This allows commands and information to be exchanged between the Arduino and the Arduino IDE as development environment.

The Arduino is connected as an agent with the sensor and actuator shields building a subsystem. The shields are arranged according to the number of pins comprising the port connection. The communication between the Arduino and the actuator shields is unidirectional. The communication between the Arduino and the sensor shields is bidirectional via a “Request/Response Communication Channel” (see Fig. 2).

4.2 Wearable Shields

In order to ensure the extensibility of WearIT, WearIT should be extensible through the shield functionality. These shields are intended to be modeled on the concepts of electronics. The specific requirements for the shields in the frame of the wearable prototyping platform WearIT are: (i) The shields should be freely placeable on the outside and inside of the vest. (ii) The shields should be quickly attachable. (iii) A simple connection should be possible between the shields and the motherboard. (iv) The number of pins of the shields should correspond to the respective application, with sensors and actuators. (v) The shields should

have good electrical properties. (vi) The deployable shields should be expandable with additional shields. (viii) New shields should be easy to manufacture.

Due to these specific requirements, various ideas for the construction of the shields were available. One idea was attaching the shields with snap buttons to the vest. The positions of the sensor and actuator shields are given by the snap buttons. At the same time, the snap buttons are also contacts, via which signals can be exchanged.

The disadvantage of the snap buttons, however, is that they require the fixing of a standard amount of pins. This excludes sensors and actuators with more than five pins. In addition, the shields are no longer freely placeable because the placement is defined by the snaps. Furthermore, the electrical properties are rather critical in this connection. Other disadvantages are that the electrical wires would have to be sewn in snaps. This would make the production of new shields more difficult.

Another idea is the connection of the shields via hook and loop fastener (Velcro) in connection with jack plug. The plug connection via jack plug has already been used successfully tried on MySignals [16]. In this approach, the cables would be pre-routed on the vest, starting on the mainboard and to the shields through the jack socket.

The problem here is that, due to the pre-routed cables, a free placement of the shields is not possible. In addition, jack plugs are limited to a maximum of five pins. This would lead to the exclusion of sensors and actuators with more than five pins. Moreover, jack sockets are difficult to mount on the shield.

To meet the requirements of the shields, the following basic structure has been defined for the shields (see Fig. 3): The first layer forms a hook and loop fastener (Velcro) surface. Since it is difficult to sew on the Velcro itself, a fabric layer is sewn on top of it. On this fabric layer, a breadboard with soldered cables is sewn on. Sensors or actuators are either soldered directly onto the board or are sewn onto the fabric and connected to the board via cables or conductive yarn. To be able to stack the shields, the shields are partially sewn with another layer of Velcro. The shields vary in size from 24×10 cm to 7×3 cm.

Table 1 lists the WearIT shields. In table, a distinction was made between the type of shield as a sensor or actuator shield. In addition, the table contains a description of the shield with the built-electronic component and the number

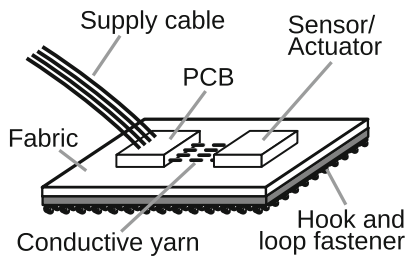


Fig. 3. Structure of the WearIT-Shield

Table 1. WearIT Shields

Type	Description	#Pins
Actuator	RGB LEDs WS2812b	4
Actuator	Vibration 2 Vib.Motors	8
Actuator	Sound Piezo-Summer	3
Sensor & Actuator	Display & Button TM1638 LED Modul	5
Sensor	Pulse SEN-11574	3
Sensor	Gyro Modul MPU9255	5

of required pins. WearIT also includes a GPS sensor. This is not available as a shield but a mainboard functionality. For this reason it is not listed in the table. Furthermore, WearIT includes free slots for generic shields. Since these do not yet have concrete functionality, these too are not listed in the table.

4.3 Mainboard

The control logic for WearIT is implemented in the mainboard. The mainboard consists of an Arduino Uno, a Raspberry Pi 3 model B, a voltage converter (5 V), a plug-in board and a GPS module. The Raspberry Pi and the Arduino are connected via a USB connection. Via the WLAN, the user can control the Raspberry Pi remotely via a VNC connection with their own computer. Thereafter, the Arduino can be programmed over the Arduino IDE on the Raspberry Pi and thus the sensor and Aktorshields be addressed. The advantage of the VNC connection is that the Arduino can be programmed without a cable connection. The Arduino can send data over a serial link to the Rapsberry Pi for further processing. A plug-in board is mounted on the Arduino via the GPIO interface. On this board there are a total of 12 slots for sensor and actuator shields (see Fig. 4). At each plug, the supply voltage VDD is present at Pin1 and the ground GND at the last pin. The 3-pin connectors are the analog signal connectors and are connected to the Arduino port A0 to A2. The 4-pin connectors are concatenated and specially designed for the WS2812b. The 5-pin connectors are I2C bus connectors and have an interrupt next to SDA and SDL. The 5-pin connectors can accept any I2C shields. The 6-pin connectors are generic input/output connectors. The connectors A1 and A2 are connected in parallel and thus share the signal lines. As a result, two shields can be inserted, which occupy only partially the existing signal lines. The connector B can switch its four input/output signal lines to SPI.

In addition, all electronic connections are on the plug-in board. The 2-pin plug is the plug for the power supply. The 5 V supply voltage is generated by a voltage converter from the battery pack.

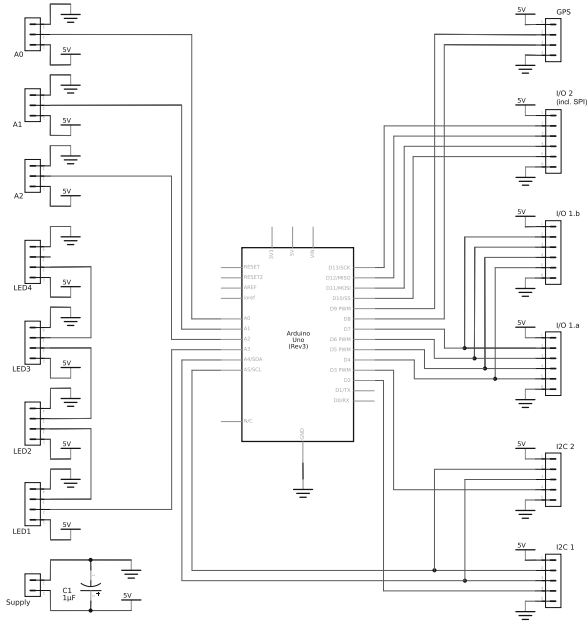


Fig. 4. Electronic schematic of the mainboard

4.4 Library

WearIT has its own library to facilitate the programming of WearIT shields. This library contains a collection of WearIT command definitions. These are intended to make it easier for the developer to access the sensor and actuator shields without specific expertise and to develop new applications based on them. The WearIT library consists of a folder “WearIT” with all existing sketches for the sensor and actuator shields. For each sensor and actuator shield, there is a header file with the extension “.h” and a C++ file. The header file contains all

```

WearIT_Buzzer_Example §
#include "WearIT_Mainboard.h"
#include "WearIT_Buzzer.h"
wearit_mainboard wearit;
wearit_buzzer buzzer;
void setup() {
    buzzer.init(wearit.a2);
}
void loop() {
    buzzer.playTone(1000);
    delay(1000);
    buzzer.stopTone();
    delay(1000);
}
    
```

Fig. 5. Example for using the library

the functions of the respective shield and the required variables. The contents of the shield functions are specified in the C++ file. In order to identify the WearIT functions as such, a uniform naming convention has been defined for the library. The naming convention includes the prefix “WearIT” and the following functions, separated by an underscore: WearIT_<function>. In Fig. 5, a buzzer program is used to demonstrate an example of using the library.

The overall using of the library is intuitive and fast, e.g. the access to the shields via the connectors is covered by the library. However, to improve the performance in large and complex programs, the library should be reworked.

5 Conclusion and Future Work

This work aimed to find a solution for a wearable prototyping platform. WearIT, a rapid prototyping platform for wearables, was developed for this purpose. This should support the rapid development of wearable prototypes. WearIT has a vest with Velcro strips, sensor and actuator, a motherboard consisting of a board, Arduino, Raspberry Pi and a GPS module and a WearIT library for programming. For the prototyping, the developer has a number of sensor and actuator shields, as well as free shields, for the use of other sensors and actuators available. To develop a prototype, the shields can be placed on the vest with Velcro and connected to the motherboard via cables. The Raspberry Pi can then be accessed via VNC and programmed using the WearIT library of the Arduino.

WearIT was able to meet all requirements at the beginning of development as well as all specific requirements during development.

Examples for using WearIT could be patient monitoring via the patient’s pulse and movement, or blind navigation in which navigation is performed using the GPS module and the vibration actuators.

Further work by WearIT would be the integration of Bluetooth for data transmission as well as possible further processing of the data. It could be possible to access the data via Wi-Fi on the data and these be further processed in an app. Also, the simplification of programming with, for example, Node-Red would also be a future work to further simplify the operation and programming with WearIT.


References

1. Mysignals: Mysignals changes the future of medical and eHealth applications. <http://www.my-signals.com/>
2. Aleksy, M., Lehtola, M., Puchakayala, A.: Design and implementation of a wearable safety suit. *IT Converg. Pract.* **1**(1), 28–36 (2013)
3. Banzl, M.: *Getting Started with Arduino*, 1st edn. Make: Projects (2008)
4. Buechley, L., Eisenberg, M.: The lilypad Arduino: toward wearable engineering for everyone. *IEEE Pervas. Comput.* **7**(2), 12–15 (2008)
5. Buechley, L., Eisenberg, M., Catchen, J., Crockett, A.: The lilypad arduino: using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. In: *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* pp. 423–432, April 2008

6. Gengnagel, C., Nagy, E., Stark, R.: Rethink! Prototyping. Transdisciplinary Concepts of Prototyping. Springer International Publishing, Cham (2016). <https://doi.org/10.1007/978-3-319-24439-6>
7. Grosse-Puppendahl, T., Berghoefer, Y., Braun, A., Wimmer, R., Kuijper, A.: Opencapsense: a rapid prototyping toolkit for pervasive interaction using capacitive sensing. In: IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 152–159 (2013)
8. Hodges, S., Taylor, S., Villar, N., Scott, J., Helmes, J.: Exploring physical prototyping techniques for functional devices using. Net Gadgeteer. In: Proceedings of 7th International Conference on Tangible, Embedded and Embodied Interaction, pp. 271–274 (2013)
9. Hung, K., Zhang, Y.T., Tai, B.: Wearable medical devices for tele-home healthcare. In: 26th Annual International Conference of IEEE Engineering in Medicine and Biology Society, IEMBS 2004, vol. 2, pp. 5384–5387. IEEE (2004)
10. Jutila, M., Rivas, H., Karhula, P., Pansar-Syväniemi, S.: Implementation of a wearable sensor vest for the safety and well-being of children. *Procedia Comput. Sci.* **32**, 888–893 (2014)
11. Lee, J.C., Avrahami, D., Hudson, S.E., Forlizzi, J., Dietz, P.H., Leigh, D.: The Calder toolkit: wired and wireless components for rapidly prototyping interactive devices. In: Proceedings of 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, pp. 167–175 (2004)
12. Oxter, J.: Arduino shield list (2017). <http://shieldlist.org/>
13. Park, S., Mackenzie, K., Jayaraman, S.: The wearable motherboard: a framework for personalized mobile information processing (PMIP). In: Proceedings of 39th Annual Design Automation Conference, pp. 170–174. ACM (2002)
14. Park, S., Jayaraman, S.: Enhancing the quality of life through wearable technology. *IEEE Eng. Med. Biol. Mag.* **22**(3), 41–48 (2003)
15. Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., Seneviratne, A.: A survey of wearable devices and challenges. *IEEE Commun. Surv. Tutor.* **19**(4), 2573–2620 (2017)
16. S.L., L.C.D: Mysignals SW. eHealth and medical IoT development platform. technical guide, July 2017. http://www.libelium.com/downloads/documentation/mysignals_technical_guide.pdf
17. VivoMetrics: Vivometrics lifeshirt system technology, March 2003. <http://www.vivometrics.com>



A Review of Health Monitoring Systems Using Sensors on Bed or Cushion

Massimo Conti¹ (✉), Simone Orcioni¹ ,
Natividad Martínez Madrid², Maksym Gaiduk³, and Ralf Seepold³

¹ Università Politecnica delle Marche, Ancona, Italy
{m.conti, s.orcioni}@univpm.it

² Reutlingen University, Reutlingen, Germany
natividad.martinez@reutlingen-university.de

³ HTWG Konstanz, Konstanz, Germany
{maksym.gaiduk, ralf.seepold}@htwg-konstanz.de

Abstract. How technology can answer the challenge that currently population ageing is facing to the healthcare system? In this work, systems and devices related to “smart” bed and cushion, that are commercially available or matter of research works, are reviewed.

Keywords: Smart bed · Smart cushion · Population ageing · Smart-care

1 Introduction

European people are living a longer and healthier life. The rapidly growing of the elderly population is happening together a reduction in fertility rates that poses great challenge to the social protection systems in Europe [1].

In this work, we are mostly interested to the challenges that the population ageing is bringing up to the health care and long-term care programs in Europe. This is also a main focus of Horizon 2020 program, from the exact words of the Commissioner for Research, Innovation and Science, European Commission, Máire Geoghegan-Quinn: “Hundreds of millions of euros have been invested over recent years in health-related ageing research, neurodegenerative diseases, including Alzheimer’s, as well as in socioeconomic dimensions of ageing. The challenge-based approach of Horizon 2020—the EU’s new programme for research and innovation—fits perfectly with the pressing need to address issues resulting from demographic change and ageing” [1]. Then she concluded how it is important to “ensure the good quality of life and well-being of the elderly” [1]. While the prevision in [2] says that in 2015, in the whole Europe the over-60 population would be the 23.9% and will increase to 29.6% in 2030, from Fig. 1 we can see that in 2014 in six countries this percentage was just over the 30%.

Daily mobile healthcare service becomes more and more important for disabled persons, people with chronic diseases and elderly people. Chronic diseases influence the health of the people living alone in daily life and a non-invasive, easy to use, real-time remote health monitoring service is required.

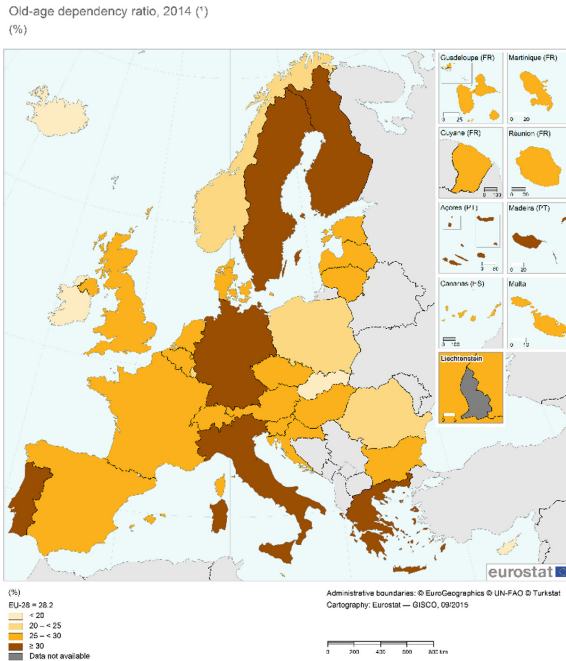


Fig. 1. Population ageing in Europe. Source: [2].

Pressure ulcers are skin injuries caused by long term exposition to high pressures on support points that interrupt blood. Pressure ulcers are an important health problem due to the enormous economical and sanitary repercussions that they imply [3].

Can the continuous progress of technology give some effective solutions for care? Does technology answer to the request for good quality of life and well-being of the elderly people and people with chronic diseases?

In this work, we will face these questions by making a review of commercial systems and components related to “smart” bed and cushion. Then a review of recently research works on the same topic will follow.

2 Commercial Devices

Commercial noninvasive monitoring systems placed on a bed or cushion are mainly of two categories:

- systems for the monitoring of pressure maps. Usually they consist in a grid of many resistive or capacitive pressure sensors. They give a detailed map of the pressure distribution of the body in the bed or chair with the aim of verifying the correct position of the person. The devices are used for example to verify the efficiency of the mattress or cushion in avoiding decubitus ulcers for the specific patient. Usually they make accurate measurements and they are suitable for designers and clinicians

to measure and visualize the key pressure points in a chair or mattress, and not for the final user. They are complex to use, require power supply, and the cost is too high to be used in the everyday life. They are used for optimizing, testing and analysis by companies, for durability and longevity studies, for seating and positioning research, for marketing tool to demonstrate the comfort of the product to the customer, and to optimize the system for the target person in the buying phase. Software tools on PC are provided to visualize the pressure map.

- systems for the monitoring of vital signals. Usually these devices have a limited number of sensors (pressure or accelerometer) for the indirect measurements of specific signals: breathing or heart rate. The accuracy is not high, the target is simple such as the presence of the persons in the bed, estimation of the sleep monitoring phases or vital signs. In many cases the target client is the final user and therefore the device is easy to use and connected to a smartphone.

Table 1 summarizes the main features of the reviewed systems, highlighting the measured quantity and the type of sensor used. The main application topics of each device are also summarized in the last column. The next subsections reports detailed information on the main commercial devices considered.

Table 1. Summary of commercial devices.

Company	Quantity	Features	Application
Novel	Pressure	Capacitive sensors, up to 1024, Bluetooth	Cushion design; wheelchair personalization
Tekscan	Pressure	Resistive, up to 16128, WiFi	Cushion and mattress design; wheelchair personalization, pressure offloading
Pressure Profile	Pressure	Capacitive, up to 512, WiFi	Cushion design
Texisense	Pressure	1024 sensors	Cushion design
Sensors Products	Pressure	Piezoresistive, up to 1728, wireless connection	Mattress selection
Emfit safebed	Pressure	Quasi-piezoelectric thin film	Bed occupancy; sleep tracking
Murata bed sensor	Acceleration	Ballistocardiography MEMS	HR, HRV, BR, and RSV measurements; bed occupancy
Early sense	Acceleration	Piezoelectric	HR, BR, Motion; fall and pressure ulcers prevention

2.1 Pressure Maps Monitoring

Novel: pliance-HS

The Novel pliance-HS system [4] provides a dynamic quantification of the sitting pressure points of patients in wheelchairs. The system helps in selecting the appropriate

cushions and correct adjustments of the wheelchair to fit the individual patient. It functions with Bluetooth wireless telemetry systems.

The system connects up to 1024 sensors to a desktop or notebook. Collected data can be stored on the internal flash memory or transmitted via Bluetooth. In spite of its high spatial resolution and advanced characteristics, it is not meant to be used for a daily use as an anti-decubitus system. Figure 2 shows a picture of the sensors part of the system.



Fig. 2. Novel pliance-HS. Source: [4].

Tekscan: Body Pressure Measurement System

It measures the pressure distribution of a human body on support surface such as wheelchairs cushion or mattresses. It provides up to 2016 sensing elements for the cushion, multiple elements can be placed to cover a mattress surface with up to 16.128 elements [5]. The target of the device is to enable precise measurement of the location and magnitude of peak pressures and overall pressure distribution patterns. The sensors are thin and flexible force sensor resistors. A similar product is applied to feet to measure pressure information for gait analysis and avoiding diabetic foot ulcers (Fig. 3).



Fig. 3. Tekscan body pressure measurement system. Source: [5].

Pressure Profile: Body Pressure Mapping for Seat and Foot

It measures and visualizes the pressure points when a user sits on a seat. Consists in an array of 512 sensors, 12 bit ADC resolution, 2 W power dissipation, WiFi wireless connection [6]. In addition, the company provides a Foot Pressure Mapping System designed for footwear or orthotic designers, researchers in gait analysis, diabetic foot complications, or other disciplines interested in the pressure distribution on the foot. It consists in 25 capacitive sensing elements with a Bluetooth wireless connection.

Texisense: TexiMat Smart cover” for a Wheelchair Cushion

The TexiMat [7] cover allows to measure the pressure map. It includes 1024 sensors and is equipped with a battery. The company provides in addition a pressure sensor sock capable of measuring a normal pressure using textile fibers (Fig. 4).

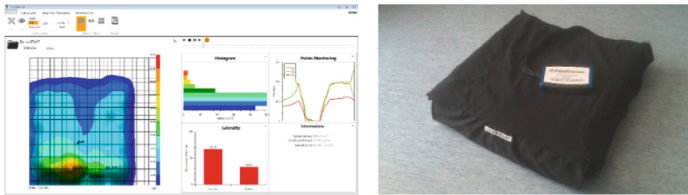


Fig. 4. Texisense: TexiMat smart. Source: [7].

Sensors Products: Pressure Monitoring System on a Mattress (Tactilus Bodyfitter)

The company Sensors Products [8] provides a pressure monitoring system on a mattress consisting in an array of 1728 piezoresistive sensors with a 35 Hz scan speed, with wireless connection. The objective is to enable the consumer to purchase a truly well fitted mattress. An additional product is a sheet for mattress for temperature measurements (Fig. 5).

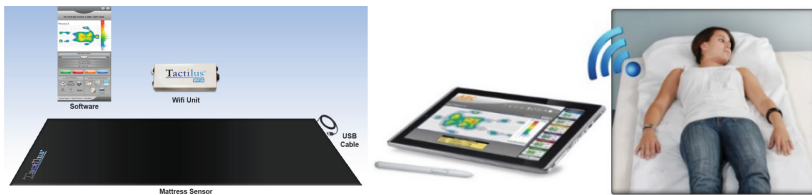


Fig. 5. Sensors products. Source: [8].

2.2 Vital Sign Monitoring

Emfit Safebed

The Emfit Safebed [9] is a bed occupancy monitoring system for fall and wandering prevention. It consists of a monitor device and an under-mattress bed sensor, it operates as a fall or bed exit monitor and it monitors the presence or absence of a person in bed by detecting all the person’s movements and micro movements, such as those caused by a person’s heart beating. The bed sensor uses a thinfilm ferro-electret technology (Electroactive Polymer) and it is placed under the mattress. The monitor has an audible notification and a dry-contact output for connection to nurse call systems or personal emergency phones. Figure 6 shows the Emfit Safebed system.



Fig. 6. The Emfit Safebed and Safebed QS. Source: [9, 10].

Emfit Safebed QS [10] is a health and wellness tracker. It tracks sleep quality and quantity, heart and breathing rates, and even heart rate variability. Changes in long-term trends can give very valuable information about health and wellbeing. It is provided in three different version, aimed respectively to improve the performances of sport amateurs or professionals, for assisted living & nursing homes, and as a clinical tool in sleep research studies. Figure 6 shows also the Emfit Safebed QS system.

Murata: Contactless Bed Sensor SCA11H

It is a device aimed to patient monitoring and elderly care in hospitals, assisted living and even at homes. The signal processing of heart and respiration rate and beat-to-beat time enables patient monitoring from bed occupancy to condition quality analysis.

The balistocardiography (BCG) technology [11] allows to output vital signs and bed occupancy information such as: heart rate, respiration rate, heart rate variability (correlates to stress level), relative stroke volume (how much blood the heart pumps), bed status indicator (is the bed empty, is patient in bed, is patient moving in bed).

The product is not meant for sale to the end user but is targeted for software solution and service providers and OEM system integrators. Figure 7 shows the Murata device.



Fig. 7. Murata contactless bed sensor. Source: [11].

Early Sense

Early sense system [12] was designed with the aim to provide continuous patient monitoring by mean of the measurements of heart rate, respiratory rate and motion of a patient in bed. These data potentially allow the clinical team to manage fall prevention, patient health state and pressure ulcers prevention.

The system has been developed specifically for medical surgical, general care patients who are usually monitored by nurses once every four to six hours. Skilled nursing facilities are emerging as important centres of care and attention for elderly people. All data came from a sensing plate with an integrated piezo-electric sensor that is placed under the mattress. The piezoelectric sensor, integrated into a membrane

plate, detects mechanical vibrations of the heart cardio ballistic (motion) effect, respiratory and patient motion. The signal is sent to a bedside monitor.

The position of the sensing plate in bed it is shown in Fig. 8.



Fig. 8. A detail of the early sense sensing plate. Source: [12].

3 Research Works

The papers reviewed in this section are recent papers on smart bed and cushion published in international journal or conferences. They present sensors, systems and methodologies to improve system performances.

Application

The application considered in papers [13–24] is the monitoring of the pressure in a cushion on wheelchairs. In many cases, the idea presented can be applied to a mattress too. The main application is the estimate of sitting posture with the aim of the decubitus ulcer avoidance. Papers [26–30] are devoted to heart and respiration rate monitoring using not invasive devices placed on the bed, while papers [31–35] estimates from the sensor data the different sleep phases.

Classification Algorithm

Great part of the work of research papers is devoted to the description and optimization of the posture classification algorithm. In some of these papers, an algorithm for posture recognition is presented, and in paper [13, 17, 19, 20] this is the main focus of the paper. An interesting state of the art of posture recognition algorithms is reported in [17] and [20].

The systems in [13, 17, 19, 20] recognize different sitting postures (15-9-7-5 positions for example: leaning left front, leaning front, leaning right front, leaning left, seated upright, leaning right, leaning left back, leaning backward, leaning right back).

The algorithms are based on: dynamic time warping [13], bayesian probability [17], supervised learning techniques [19]. In [20] the classification algorithms are: decision stump (a learning model that is a one-level decision tree), naive Bayes (a parametric classifier based on bayesian theorem with assumption that all features are independent), and AdaBoost C4.5 (a decision tree algorithm which is commonly used as a weak classifier in AdaBoost). The classification accuracy declared in [13, 17, 19, 20] are high enough (over 80%).

The estimation of the position in [21] is based on the variation of the temperature over time. No indication on classification algorithm and accuracy are given.

Table 3 reports some specific information reported in the paper on the type and number of pressure sensors used. In general, the researches presented use a reduced number of sensors. Only paper [31] uses a high-resolution device, but in this case the device is a commercial device and the focus of the paper is the sleep classification algorithm.

Table 3. Sensor and actuators types of the research works in references.

	No. of sensors	Pressure sensor type	Pump type	Electro-valves type
[13]	16 × 16	Eeonyx		
[14]	10 × 10	AdvanPro		
[16]	12	Freescale	Airpro	Festo
[17]	2 × 8 + 8 × 8	(binary sensor)		
[18]	4	Keyence	Okenseiko	TDL
[19]	8	Vista medical		
[20]	5 × 8	Flexiforce by Tekscan		
[22]		Precision seating solutions		
[23]	8	FSR-406 by interlink electronics		
[24]	4	MPX53GP by Freescale		
[31]	64 × 27	Vista medical		

The paper [13] proposes a cushion, aimed to estimate the sitting postures. It is based on a textile resistive piezoelectric sensor array, a 16 × 16 array of 1 cm × 1 cm sensors with a spacing of 50 mm.

In [14] the authors present a new method for the readout of resistive sensor arrays compared to transistor/diode controlled approach, multiplexer and op-amp assisted approach and incidence matrix approach. The advantages are reduced circuitual complexity, improved accuracy and low crosstalk error. The authors implemented a cushion prototype of 10 × 10 textile resistive sensor array for sitting pressure monitoring.

The paper [15] presents a home mobile healthcare system for wheelchair users, using IoT technologies. The paper focus on the app on the smartphone that collects the data from the sensors (heart rate and ECG sensors, pressure detecting cushion), and on the web service.

The work presented in [16] describes a measurement system of pressure distribution on a chair. It consists of a cushion with 12 air-cells, 12 air pressure sensors, 12 micro electro valves, one air micro pump. The main advantage are the 12 independent air cells with respect to the commercial alternate mechanism.

The paper in [17] presents a sensing cushion with 2 × 8 + 8 × 8 binary sensors, scanned with a 10 Hz frequency by a micro controller unit with a Bluetooth module.

The paper [18] presents depressurization control scheme using pressure distribution estimation. The system consists of a seating cushion with 4 air cells which can lift or incline the seating cushion. Each air cell has a pressure sensor, a solenoid valve and a small air compressor.

The system [19] utilizes piezo resistive sensors beneath a wheelchair cushion to monitor pressure. The system consists of 8 pressure sensors (Vista Medical).

The authors of paper [20] implemented a sitting posture recognition system, which can recognize sitting posture. The cushion consists in a pressure sensor array: 5×8 low cost sensors (15\$, Tekscan Flexiforce) and a 16-bit ADC.

The paper [21] presents a posture recognition system based on an array of temperature sensors. Sensors are placed on the cushion (4×3) and on the two armrests (3 each). The estimation of the position is based on the variation of the temperature over time. A pressure sensor for wheelchair (Sensimat) was used to test the system.

In [22] an app was developed to track, using the smartphone accelerometer, the movements of the patient and to monitor if the user is awake or asleep.

In paper [23] a sensing cushion has been developed to determine the user's posture. 8 pressure sensors are placed on the cushion (Force Sensing Resistor FSR-406 by Interlink Electronics has been used).

The paper [24] presents a wheelchair cushion pressure monitoring system, designed to detect changes in air pressure of a wheelchair cushion. The system consists of a thin pad with 4 load cells, placed under the cushion and connected by wire to the monitoring system. The system [24] is not used for posture classification. Alert is delivered by the system via BLE to a smartphone in case of low air pressure.

The paper [25] evaluates the effect on interface pressure of a proposed alternating pressure sequence for an air-cell cushion. The cushion consists of 12 air cells interconnected in pairs. Each cell has an air pressure sensor and an electrovalve. Estimation of mean value of interface pressure and contact area of foam cushion, static cushion and alternating cushion.

The papers [26–29] present a pneumatic biomeasurement system for monitoring the heartbeat and respiration. Heartbeat and respiration microvibrations cause changes in air pressure of a thin air cushion placed between the bed frame and the coil-spring cushion. The vibrations are detected by a pressure sensor. An analytical model of the pneumatic filter is developed.

In [30] a pressure sensor grid is placed between the mattress and the bed frame for respiration rate, movement and sleep monitoring.

In [31] a classification approach of 3 sleep states (sleep, pre-wake, wake) is proposed that uses surface pressure sensors. The pressure system (Vista Medical with 64×27 sensors) is a flexible and thin layer placed on top of the bed mattress.

In [32–35] different classification algorithms of 2 sleep states (sleep, wake) are proposed with accuracy between 84% to 95%. The authors use respiration and actigraphy sensors [32], respiration and ECG sensors [33], ECG sensors [34], actigraphy sensors [35].

4 Conclusions

The commercial mattress and wheelchair cushion used to avoid decubitus ulcers have not the ability to estimate the posture of the user. Commercial devices consisting of matrix of pressure sensors, such as the Novel pliance-HS system, may have a great number of sensors and they can be used for research purpose for the posture estimation, but the high cost of the device does not allow the use in commercial mattress or cushion.

A smart device that with a limited number of sensors is able to adapt the pressure of the single cell of the cushion or mattress to the actual posture of the user does not exist.

Many of the research papers try to improve posture recognition, heart rate or respiratory rate estimation using signal processing of the sensor data and recognition algorithms. A future research and commercial step could be the development of a smart device that with a limited number of sensors able to adapt the pressure of the single cell of the cushion or mattress to the actual posture of the user.

References

1. European Commission: Population ageing in Europe, Facts, implications and policies, Luxembourg: Publications Office of the European Union (2014)
2. European Union (EU): People in the EU – Population Projections, June 2015. <http://ec.europa.eu/eurostat/statistics-explained/>
3. European Pressure Ulcer Advisory Panel. <http://www.epuap.org/pu-guidelines/>
4. Novel. <http://www.novel.de/novelcontent/pliance/wheelchair>
5. Tekscan. <https://www.tekscan.com/>
6. Pressure Profile. <https://pressureprofile.com/>
7. Taxisense. http://www.taxisense.com/capteur_en
8. Sensors Products. <http://www.sensorprod.com/dynamic/mattress.php>
9. Emfit. <http://www.safetyssystemsdistribution.co.uk/bed-exit-alarms/>
10. Emfit. <https://www.emfitqs.com/>
11. Murata. <https://www.murata.com/>
12. EarlySense. <https://www.earlysense.com/>
13. Xu, W., Huang, M., Amini, N., He, L., Sarrafzadeh, M.: eCushion: a textile pressure sensor array design and calibration for sitting posture analysis. *IEEE Sens. J.* **13**(10), 3926–3934 (2013)
14. Shu, L., Tao, X., Feng, D.D.: A new approach for readout of resistive sensor arrays for wearable electronic applications. *IEEE Sens. J.* **15**(1), 442–452 (2015)
15. Yang, L., Ge, Y., Li, W., Rao, W., Shen, W.: A home mobile healthcare system for wheelchair users. In: *Proceedings of IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 609–614 (2014)
16. Arias, S., Cardiel, E., Garay, L., Tovar, B., Pla, M., Rogeli, P.: A pressure distribution measurement system for supporting areas of wheelchair users. In: *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4751–4754 (2013)
17. Xu, L., Chen, G., Wang, J., Shen, R., Zhao, S., A sensing cushion using simple pressure distribution sensors. In: *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 451–456 (2012)
18. Chugo, D., Fujita, K., Sakaida, Y., Yokota, S., Hashimoto, H.: A depressurization motion assistance for a seated wheelchair user using pressure distribution estimation. In: *5th International Conference on Human System Interactions*, pp. 75–80 (2012)
19. Dai, R., Sonenblum, S.E., Sprigle, S.: A robust wheelchair pressure relief monitoring system. In: *Annual International Conference of IEEE Engineering in Medicine and Biology Society*, pp. 6107–6110 (2012)
20. Liang, G., Cao, J., Liu, X.: Smart cushion: a practical system for fine-grained sitting posture recognition. In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 419–424 (2017)

21. Russell, L., Goubran, R., Kwamena, F.: Posture sensing using a low-cost temperature sensor array. In: *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 443–447 (2017)
22. Khan, A., Reuter, M., Phung, N., Hafeez, S.: Wireless solution to prevent decubitus ulcers: preventive weight shifting guide, monitor, and tracker app for wheel chair users with spinal cord injuries (phase II). In: *IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6 (2016)
23. Ahn, B., Kim, T., Sain, M., Jeong, D.: Implemented of posture discriminate system through the seat weight distribution. In: *IEEE Conf. on Wireless Sensors (ICWiSe)*, pp. 24–27 (2015)
24. Sazonov, E., Abele, R., Gerrity, A., May, P., Spradling, K., Kennamore, T., Chen, Y., Klebine, P.: Development of wheelchair cushion pressure monitoring system. In: *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–5 (2014)
25. Arias, S., Cardiel, E., Rogeli, P., Mori, T., Nakagami, G., Noguchi, H., Sanada, H.: An alternating pressure sequence proposal for an air-cell cushion for preventing pressure ulcers. In: *Proceedings of Conference on IEEE Engineering in Medicine and Biology Society*, pp. 3480–3483 (2014)
26. Kurihara, Y., Watanabe, K.: Suppression of artifacts in biomeasurement system by pneumatic filtering. *IEEE Sens. J.* **12**(3), 416–422 (2012)
27. Bao, J., Shou, X., Wang, H., Yang, H.: Study on heartbeat information acquired from pressure cushion based on body sensor network. In: *IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pp. 1103–1108 (2013)
28. Watanabe, K., Watanabe, T., Watanabe, H., Ando, H., Ishikawa, T., Kobayashi, K.: Noninvasive measurement of heartbeat, respiration, snoring and body movements of a subject in bed via a pneumatic method. *IEEE Trans. Biomed. Eng.* **52**(12), 2100–2107 (2005)
29. Tanaka, S., Matsumoto, Y., Wakimoto, K.: Unconstrained and non-invasive measurement of heart-beat and respiration periods using a phonocardiographic sensor. *Med. Biomed. Eng. Comput.* **40**(2), 246–252 (2002)
30. Gaiduk, M., Kuhn, I., Seepold, R., Ortega, J.A., Madrid, N.M.: A sensor grid for pressure and movement detection supporting sleep phase analysis. In: Rojas, I., Ortuño, F. (eds.) *IWBIO 2017. LNCS*, vol. 10209, pp. 596–607. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56154-7_53
31. Baran Pouyan, M., Nourani, M., Pompeo, M.: Sleep state classification using pressure sensor mats. In: *37th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, 26 August 2015, pp. 1207–1210 (2015). ISSN 1094-687X
32. Long, X., Fonseca, P., Foussier, J., Haakma, R., Aarts, R.: Sleep and wake classification with actigraphy and respiratory effort using dynamic warping. *IEEE J. Biomed. Health Inf.* **18**, 1272–1284 (2013)
33. Karlen, W., Mattiussi, C., Floreano, D.: Sleep and wake classification with ecg and respiratory effort signals. *IEEE Trans. Biomed. Circuits Syst.* **3**(2), 71–78 (2009)
34. Lewicke, A., Sazonov, E., Corwin, M.J., Neuman, M., Schuckers, S.: Sleep versus wake classification from heart rate variability using computational intelligence: consideration of rejection in classification models. *IEEE Trans. Biomed. Eng.* **55**(1), 108–118 (2008)
35. Lotjonen, J., Korhonen, I., Hirvonen, K., Eskelinen, S., Myllymaki, M., Partinen, M.: Automatic sleep-wake and nap analysis with a new wrist worn online activity monitoring device vivago WristCare. *Sleep* **26**(1), 86–90 (2003)



Textile Sensor Platform (TSP) - Development of a Textile Real-Time Electrocardiogram

Thomas Walzer¹(✉), Christian Thies¹, Klaus Meier²,
and Natividad Martínez Madrid¹

¹ School of Informatics, Reutlingen University, Reutlingen, Germany
{thomas.walzer,christian.thies,
natividad.martinez}@reutlingen-university.de

² School of Textiles & Design, Reutlingen University, Reutlingen, Germany
klaus.meier@reutlingen-university.de

<http://inf.reutlingen-university.de>, <http://td.reutlingen-university.de>

Abstract. Being able to monitor the heart activity of patients during their daily life in a reliable, comfortable and affordable way is one main goal of the personalized medicine. Current wearable solutions lack either on the wearing comfort, the quality and type of the data provided or the price of the device. This paper shows the development of a Textile Sensor Platform (TSP) in the form of an electrocardiogram (ECG)-measuring T-shirt that is able to transmit the ECG signal to a smartphone. The development process includes the selection of the materials, the design of the textile electrodes taking into consideration their electrical characteristics and ergonomics, the integration of the electrodes on the garment and their connection with the embedded electronic part. The TSP is able to transmit a real-time streaming of the ECG-signal to an Android smartphone through Bluetooth Low Energy (BLE). Initial results show a good electrical quality in the textile electrodes and promising results in the capture and transmission of the ECG-signal. This is still a work-in-progress and it is the result of an interdisciplinary master project between the School of Informatics and the School of Textiles & Design of the Reutlingen University.

1 Introduction

Monitoring heart activity is an important diagnostic method to assess a person's individual health status and fitness level. Therefore electro cardiograms (ECG) outperform devices using photoplethysmograms (PPG) such as consumer fitness wearables, since they do not only measure heart rates. The pattern of the electric stimulation provides precise information on heart rate variability (HRV) or cardiac disorders. However, the measurement of the electric signal requires close contact of the classical wet electrodes to the skin, which have to be applied for each examination by trained personal. During holter monitoring a person is

The original version of this chapter was revised: The acknowledgement was added. The correction to this chapter is available at https://doi.org/10.1007/978-3-319-78759-6_42

limited in body hygiene and must consider the cables and handle the recording device. Overall, long-term ECG data acquisition becomes inconvenient, making it an exceptional diagnostic tool not suitable for everyday life. But long-term or spontaneous monitoring of heart activity is important for applications such as workout, health assessment or personalized medicine where an easy to apply method is required [1]. Therefore current approaches to integrate ECG sensors into more practical wearable forms such as clothing include unobtrusive and automatic acquisition that makes use of dry electrodes [4]. However these plastic electrodes require a large surface that is tightly pressed to the body and that causes sweating. For that reason and since there is not yet enough medical evidence available, this new approach has so far been applied mainly into fitness outfit such as the *ambiotex shirt*¹. There is some evidence that individual long-term monitoring supports personalized medicine in general, but clinical trials supporting health care decision makers need more systematic approaches and tools². For that purpose, a new sensor shirt based on conductive filaments is developed in this interdisciplinary project, which can be used like a normal piece of garment without any stiff or sticky plastic components. The integrated approach consists of the manufacturing of woven electrodes, seamless integration in the textile, individual medical data acquisition and processing, including diagnostic support, which is evaluated with respect to reliability and stability of ECG measures.

2 Related Work

The topic of personalized medicine is becoming increasingly important and, together with telemedicine, it could be a key factor for better health care, e.g. in rural areas. For personalized medicine, it is important to obtain vital data from the user or patient. Wearables are used for this purpose, whereas fitness oriented people wear them for self-optimization. An essential feature of wearables is the wearing comfort. They are mostly made of plastic, which is not pleasant on the skin and increases sweating of the wearer. There is therefore an advantage with intelligent textiles, which are the textile version of the wearable, because they can feel as a second skin to the wearer. This section describes the main components of smart textiles, introduces the basics of the electrocardiogram, describes the possibilities of Bluetooth as a wireless communication protocol between the smart textile and the smartphone and shows the differences between the ECG T-shirt under development and other wearables oriented to fitness tracking.

2.1 Smart Textiles

Smart textiles, as defined in DIN CEN/TR 16298 [3], are textiles that have a smart function [9] e.g. changing the color on contact with water or, like in

¹ <https://www.ambiotex.com/>.

² <http://med.stanford.edu/news/all-news/2017/01/wearable-sensors-can-tell-when-you-are-getting-sick.html>.

our case, measuring the ECG. Such systems could have functionality similar or identical to wearable electronics, like a fitness tracker, but integrated in textiles. The following components are important for an electronic smart textile (see also Fig. 1):

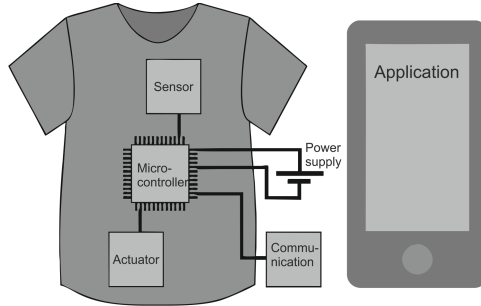


Fig. 1. A model of a smart textile T-shirt. The components microcontroller, actuator, sensor, power supply and conduction are integrated in the shirt. There is also a communication module for a connection to the smartphone.

Microcontroller The microcontroller is the component performing the data processing in the smart textile. The inputs and outputs coming from other components are connected to this device by wires or preferably by conductive yarns.

Sensors The sensors transform environment values to analog or digital values, e.g. the level of light that can be processed by the microcontroller.

Actuators The actuators transform analog or digital values from the output of the microcontroller to interactions with the environment, e.g. an LED or a speaker.

Conduction The conduction connects all the components in one or more circuits. In a smart textile, it should be a conductive yarn or fabric.

Power source The power source is typically an internal battery that can be charged or changed. An external power source is a plug direct in a power supply.

Communication There could be a communication module with a wired or wireless connection, for example a Bluetooth or wireless LAN connection.

2.2 Electrocardiogram

The electrocardiogram is a [6] method to record the electrical activity of the heart. The ECG is recorded by using leads. A lead refers to an imaginary line between two ECG electrodes. The electrical activity of this lead is measured and recorded as part of the ECG. A typical simplified configuration is the Einthoven triangle [8] with three electrodes to the body. The connection points are the

left, right arm and left leg. The sensors can be moved slightly. A key feature for recording an ECG signal is the sampling rate. While there is no significant difference between 500 Hz and 250 Hz [7], with a sampling rate of 125 Hz there is a significant quality reduction, yet it is high enough to get the ECG parameters. A sample of an ECG signal is shown in Fig. 7. The doctor can recognize the heart rate and abnormalities of the heart in an ECG.

2.3 Bluetooth Low Energy

Bluetooth [2] is a wireless communication protocol operating in the 2.4 GHz band. It is intended for short-distance connections between devices. In a smartphone, different versions of the Bluetooth protocol are integrated and implemented. Most smartphones on the market have Bluetooth Low Energy (BLE) integrated [5], which has less power consumption than a normal Bluetooth connection. The speed of Bluetooth Low Energy is up to 1 Mbps and has a range of up to 10 m. The next version of Bluetooth is number 5, which can reach speeds up to 2 Mbps and a range of up to 100 m. Bluetooth Low Energy was chosen for this development because the smartphone should act as a gateway from the smart textile to the exterior, so the user could bring his or her own device and just install the corresponding app.

2.4 Fitness Tracker

There are a lot of fitness trackers on the market, typically integrated into wrists or chest straps. They are explicitly made for fitness purposes, i.e. for recreational use and without certification a medical product. A fitness product that comes close to the TSP is the T-Shirt from Ambiotex³. It has two rectangular plastic contact surfaces to the body with a size of 70 × 30 mm. On the front of the T-Shirt it is placed the “TechUnit”, which collects the data and transmits them via BLE. It is attached to the shirt with 4 magnetic fasteners. When connected to the app, the heart rate and RR interval are displayed. This is also possible without the Ambiotex app via a BLE connection. Ambiotex uses the data to determine for example the user’s stress level. In order to establish a good connection between the body and the electrodes, there is a tighter strip below the chest. The two electrodes are integrated on this strip. In a self-experiment, however, this was perceived as uncomfortable, as the pressure to the body was quite high. It was also hard to feel if the “TechUnit” was still on the body when there were several layers of clothing (shirt and jacket) over the T-shirt.

3 Context

The development of this Textile Sensor Platform (TSP) in the form of a smart T-shirt able to measure the ECG is part of a larger project for personalized

³ <https://www.ambiotex.com/>.

medicine aimed at providing the entire pipeline from the patient over the smartphone to the clinical system. The pipeline is shown in Fig. 2. The left part shows the phase of capturing the user's ECG data through the smart T-shirt. This is the Textile Sensor Platform (TSP) development shown in this paper. The middle part of the figure shows the smartphone as a gateway for receiving data from the ECG shirt. Here, the data is preprocessed and transmitted to the clinical system over the Internet. The final step on the right of the figure shows the clinical system, which retrieves the information and stores it to enable the physician to access the data. The whole pipeline is being developed at Reutlingen University.

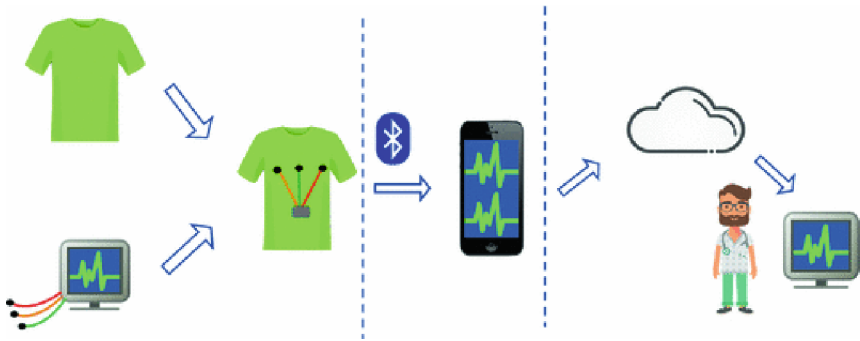


Fig. 2. System context. The ECG signal is captured with the T-shirt and sent via the smartphone to a clinical system.

For the interconnection between the T-shirt and the smartphone, an ECG profile for Bluetooth Low Energy has been defined together with some middleware to generate the communication stubs for the microcontroller on the T-Shirt and the smartphone. Another development has centered on studying the possibilities and limitations of the usage of BLE for the streaming of signals.

The goal of the Textile Sensor Platform (TSP) project is to develop a wearable ECG system as an interdisciplinary project between the School of Informatics and the School of Textiles & Design of Reutlingen University. It should be a deeply integrated ECG in smart textile. The first part of the design and development was devoted to experiment and obtain a textile prototype with textile sensors (electrodes) and connections able to provide the biosignal input to the ECG with adequate quality. The development and results of this first part are shown in this paper. In the second part of the TSP project, the ECG signal should be further processed in the microcontroller. One goal is to try and minimize artifacts, specially due to movement. Another goal is to extract relevant information out the signal, like the RR-interval, thus minimizing under certain circumstances the data to be transmitted to the smartphone and further on to the clinical system.

4 Requirements

The goal of the TSP is to integrate an ECG system into a smart textile. The components of a clinical ECG system should be integrated into a T-shirt. Instead of self-adhesive plastic pads, the electrodes should be made of textile yarns or fabrics. Furthermore, the conduction should be made of textile fabric or yarn instead of a shielded cable. The experimental model or prototype should have the following requirements. It should be worn as a normal T-shirt. It should work when the patient connects microelectronics in the T-shirt with a smartphone. Finally, it should measure the ECG signal on skin contacts.

The advantage of an ECG that is fully integrated in a T-shirt is that it can be fully customized to the user. The ideal fit can be achieved by measuring the user's body.

5 Development

The development of the ECG T-shirt was performed in four phases. The first phase was dedicated to the research on smart textiles, ECG principles and conductive material. In the second phase, different possibilities for the textile electrodes were built and tested. In the third phase, the selected version of the electrodes was integrated in the T-shirt and connected through conductive yarn. Finally, the first version of the electronic was integrated in the prototype in order to analyze the quality of the obtained signals.

5.1 Selection of Textile, ECG Design and Conductive Material

The smart textile was selected to be a T-shirt, because it stretches over the entire upper body and an ECG based on the Einthoven lead could be realized there. The electrodes should be made of conductive yarn. The connections should be flexible and have the same length.

The electronic resistance of the different conductive materials was measured. Two types of yarn were used in the prototype. One yarn was tinned copper wire and the other a twisted yarn combining a silver coated polyamid (PA) yarn and a polyester (PET) yarn.

5.2 Design of the Electrodes

The electrodes are the connection point between the body and the ECG system. The shape and material of the electrodes determine the quality of the ECG signal. They must sit well on the skin but should not be felt or noticed. The three most relevant questions for the design of textile electrodes are: (1) which structure and shape are best to acquire a good signal, (2) how electrodes can be designed to be comfortable, and (3) how to craft them.

The shape should be similar to conventional electrodes, i.e. they are round. Three different method were tested for the design of the electrodes:

sewing/stitching (embroidered), weaving and knitting. The resulting electrodes were first tested on the skin on the hands. Furthermore, they were tested on the upper body. For this purpose, the electrodes were attached to a piece of fabric and held against the upper body. In addition, the resistance of all electrodes was measured.

The self-tests of the electrodes on the skin have shown that the woven and embroidered electrodes were better placed on the skin. Further experiments revealed that the electrodes on the surface are scratchy. With the weaving method, the electrodes could be given a slight curvature with a lining. This woven electrode is more comfortable and has a better contact to the skin due to the curvature as it is pushed in a little bit. This effect would have been more difficult to achieve with the embroidered variant. Therefore, the woven electrodes were chosen. The production of these electrodes is described below.

The woven electrode is made with two types of conductive yarns. The construction is shown in Fig. 3. For the warp yarn (solid line) the twisted yarn combining a silver coated PA yarn and a PET yarn was used and for the weft yarn (dotted line) the tinned copper wire was used. The diameter of the electrodes is 3 cm. The electrode was made with a carton template and the warp yarns were fixed at the carton. The weaving started in the middle.

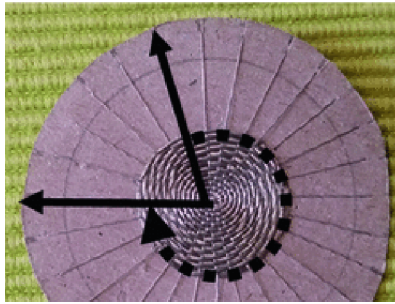


Fig. 3. Construction of the woven electrode. The radial solid arrows show the warp yarn and the circular dotted arrow the weft yarn.

For the measurement of the resistance of the electrodes, a reference signal from a signal generator (Fig. 4) was applied to the electrode and then taken off again with an oscilloscope. The parameters of the generated signal are shown in Table 1.

Table 1. Parameters of the generated signal.

Amp.: ± 20 mV	Freq.: 100 Hz	Filter: Lowpass 1 kHz	x-axis: 0–50 ms	y-axis: ± 50 mV
-------------------	---------------	-----------------------	-----------------	---------------------

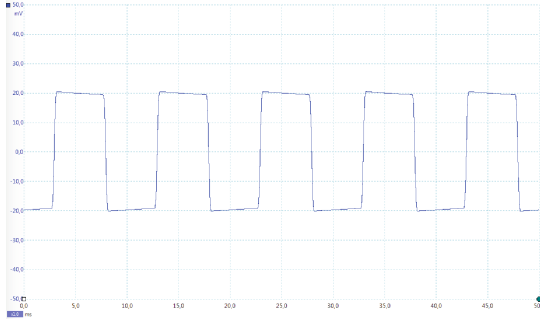


Fig. 4. Generated reference signal for the comparison of the electrodes.

The measurements of the signal on the sample of the woven electrode can be seen in Table 2. The frequency deviation is minimal and the maximum and minimum frequency is identical. The AC RMS of the voltage is slightly below the 20 mV but the minimum and maximum have a standard deviation of only 2.527 μV . The values show that the lined woven electrodes have sufficient quality for the prototype.

Table 2. Signal parameters of the woven electrode.

Frequency				AC RMS			
Min (Hz)	Max (Hz)	Median (Hz)	SD (mHz)	Min (mV)	Max (mV)	Median (mV)	SD (μV)
100	100	100	2,796	19,37	19,38	19,37	2,527

5.3 Prototype

Three experimental models (T-shirt prototypes) were created: (1) one with an attached microcontroller (without function) and a bag in front for a battery (see Fig. 5), (2) one with longer wires to connect any device for ECG measuring and (3) one that has two tissue layers, having one cover layer above the electronic.

All prototypes use the tinned copper wire as conductive yarn with a conductor path in waveform, as shown in Fig. 5 [right]. The waveform makes the wire flexible for the movements of the shirt when putting on. The waves have a distance of 1.5 cm and an amplitude of 3 cm.

The first step in the experimental models was to make the circuit path. The circuit and the electrode were made with one piece of yarn. For this, the conductive material was sewn on the front of the T-shirt. The rest of the thread was changed on the back. The electrodes were woven with the yarn using a template. After the electrode was woven, it was sewn on the front with a zig-zag stitch on the t-shirt. It fixes the electrode on the T-shirt and improves the connection.

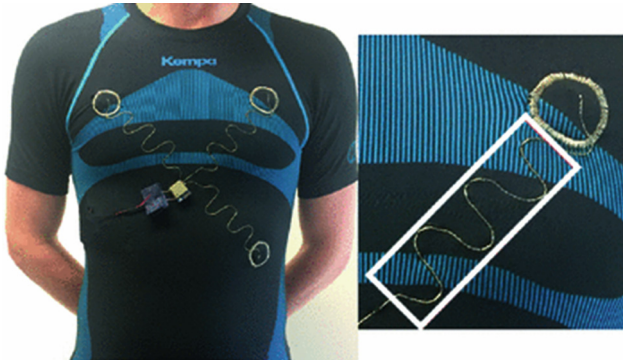


Fig. 5. Prototype with attached microcontroller (left). Zoom of the outside of the electrode and conduction (right). The white box highlights the conductive yarn in waveform.

5.4 Microelectronic

For the first prototyping phase of this project, off-the-shelf microelectronic components were used. It is a combination of two devices (shown in Fig. 6) with dimensions around 35×26 mm. One of these components is a TinyTILE⁴. It is an Arduino-compatible microcontroller based on a Nordic Semiconductor NRF51 chip with integrated BLE. The TinyTILE is a small form factor version of the Arduino 101, which was the advantage of using this microcontroller instead of the Arduino 101 for better integration. The second component is an ECG board from Spark Fun. It has 3 pins for connection to ECG electrodes. The ECG board of SparkFun⁵ has three output signals: two signals for lead-off and an analog signal with ECG information. The Arduino sketch for the ECG signal is more or



Fig. 6. Microelectronic for testing the prototype. Left a TinyTILE and right an ECG-board from Sparkfun.

⁴ <https://www.element14.com/community/docs/DOC-82913/1/tinytile-intel-curie-based-miniaturised-adaptation-of-the-arduino-genuino-101-board>.

⁵ <https://learn.sparkfun.com/tutorials/ad8232-heart-rate-monitor-hookup-guide>.

less simple to read the analog signal. The idea for the project progress is to use a NRF52 microcontroller and a real-time operating system (RTOS) on the microcontroller.

6 Results

The experimental model was tested with the microelectronics described in the previous section. In order to get a signal from the prototype, it was necessary to use an electrode gel or water. The signal obtained by using the prototype, the gel and a serial connection to the computer strongly resembles an ECG signal (see Fig. 7). The test person was in a rest position without movement. If there is a movement, e.g. with the arms, motion artifacts are displayed. During the arm movement, a slight movement of the electrodes with the T-shirt can be seen. There should also be an improvement in the fitting of the electrodes and the T-shirt.

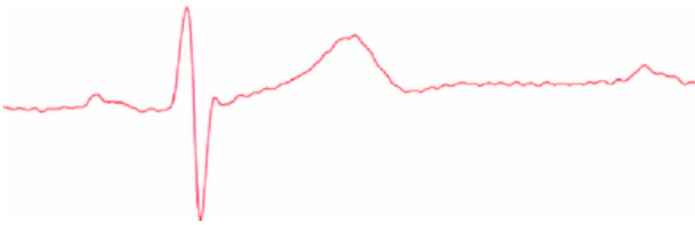


Fig. 7. Signal captured on a computer via serial connection to the microcontroller.

For the interconnection between the microcontroller and a smartphone to transmit the ECG signal, an ECG profile for Bluetooth Low Energy has been defined together with some middleware to generate the communication stubs for the microcontroller on the T-Shirt and the smartphone. A real-time streaming with the Arduino 101 and an Android application based on this BLE ECG profile was developed. An example transmission can be seen in Fig. 8. At the moment

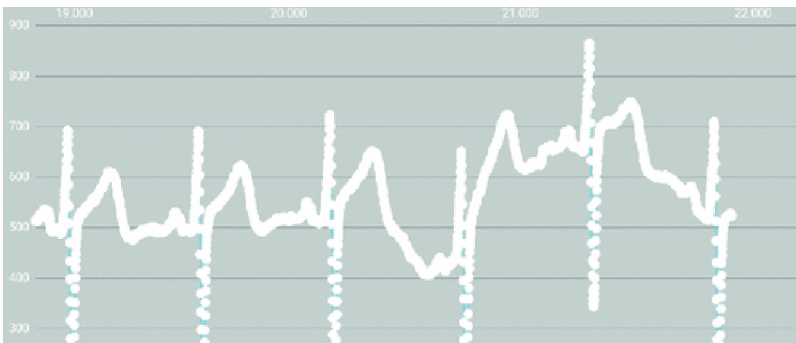


Fig. 8. Signal transmitted with BLE to a smartphone.

the visualization of the ECG signal is not optimal, but an ECG signal could be detected with a zoom. There were no strict tests for the uptime, but a simple test with a running time of 30 min.

7 Conclusion

The goal of the TSP project was to create a textile sensor platform as an interdisciplinary project between the Schools of Informatics and the School of Textiles & Design. This textile sensor platform is an ECG-measuring T-shirt. The first phase of this project has provided a functioning textile prototype. The work has started with the selection of the garment, selection of conductive yarns and ECG design. A crucial part of the project was the design of the textile electrodes. The experiments have taken into account shape, comfort, conductivity and crafting complexity. A lined woven electrode has been selected as most appropriate and its electrical properties have been measured with a reference signal before integrating the electrodes in the textile prototype. The experimental model shown in this paper is completed with an first, off-the shelf electronic circuitry connected through waveformed conductive yarn. Initial experiments show a good quality in the ECG signal measured directly on the T-shirt. The sensor platform is also able to transmit data via BLE to a smartphone. This work has shown that it is feasible to design a textile-based, ECG T-shirt. Through improvements and automated manufacturing, this sensor platform can be used as a tool for long-term medical monitoring of patients.

The next step is to continue to perform tests with the system. One important aspect is the battery life of the microcontroller and the smartphone. A change of the microcontroller architecture is also planned. With regard to the sensor platform, the electrodes should be modified in such a way that no gel or water is needed to improve the skin contact. The fitting of the T-shirt to the body should also be improved. This could reduce artifacts during migration. A common problem is the contacting of the microelectronics with the sensor platform, the connection should be very strong to have an electrical connection, but also removable for washing processes.

Acknowledgements. We thank our students Marija Elbl, Alice Fensch, Lena Rademacher, Johannes Schuh and Clemens Weißenberg for their contributions to the research presented in this paper.




References

1. Appelboom, G., Camacho, E., Abraham, M., Bruce, S., Dumont, E., Zacharia, B., Connolly, E.: Smart wearable body sensors for patient self-assessment and monitoring. *Arch. Public Health* **72**(1) (2014). <https://doi.org/10.1186/2049-3258-72-28>
2. Bluetooth SIG: Bluetooth core specification version 4.0. Specification of the Bluetooth System (2010)
3. DIN: Textiles and textile products smart textiles definitions, categorisation, applications and standardisation needs (2011)

4. Gargiulo, G., Bifulco, P., McEwan, A., Tehrani, J.N., Calvo1, R.A., Romano, M., Ruffo, M., Shephard, R., Cesarelli, M., Jin, C., Mohamed, A., van Schaik, A.: Dry electrode bio-potential recordings. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, pp. 6493–6496, August 2010
5. Gomez, C., Oller, J., Paradells, J.: Overview and evaluation of bluetooth low energy: an emerging low-power wireless technology. *Sensors* **12**(9), 11734–11753 (2012)
6. Gupta, R., Mitra, M., Bera, J.: ECG Acquisition and Automated Remote Processing. Springer, Heidelberg (2014). <https://doi.org/10.1007/978-81-322-1557-8>
7. Pizzuti, G., Cifaldi, S., Nolfè, G.: Digital sampling rate and ECG analysis. *J. Biomed. Eng.* **7**(3), 247–250 (1985)
8. Rautaharju, P., Rautaharju, F.: Investigative Electrocardiography in Epidemiological Studies and Clinical Trials. Springer Science & Business Media, Heidelberg (2007). <https://doi.org/10.1007/978-1-84628-481-6>
9. Tao, X.: Handbook of Smart Textiles. Springer Singapore, Singapore (2015)



Sensor-Mesh-Based System with Application on Sleep Study

Maksym Gaiduk¹(✉) , Bruno Vunderl¹, Ralf Seepold¹,
Juan Antonio Ortega² , and Thomas Penzel³ 

¹ HTWG Konstanz, Alfred-Wachtel-Strasse 8, 78462 Konstanz, Germany
{maksym.gaiduk, ralf.seepold}@htwg-konstanz.de,
bruno.vunderl@gmail.com

² Universidad de Sevilla, Avda. Reina Mercedes s/n, 41012 Seville, Spain
jortega@us.es

³ Sleep Medicine Center of Charité, Charitéplatz 1, 10117 Berlin, Germany
thomas.penzel@charite.de

Abstract. The process of restoring our body and brain from fatigue is directly depending on the quality of sleep. It can be determined from the report of the sleep study results. Classification of sleep stages is the first step of this study and this includes the measurement of biovital data and its further processing.

In this work, the sleep analysis system is based on a hardware sensor net, namely a grid of 24 pressure sensors, supporting sleep phase recognition. In comparison to the leading standard, which is polysomnography, the proposed approach is a non-invasive system. It recognises respiration and body movement with only one type of low-cost pressure sensors forming a mesh architecture. The nodes implement as a series of pressure sensors connected to a low-power and performant microcontroller. All nodes are connected via a system wide bus with address arbitration. The embedded processor is the mesh network endpoint that enables network configuration, storing and pre-processing of the data, external data access and visualization.

The system was tested by executing experiments recording the sleep of different healthy young subjects. The results obtained have indicated the potential to detect breathing rate and body movement. A major difference of this system in comparison to other approaches is the innovative way to place the sensors under the mattress. This characteristic facilitates the continuous using of the system without any influence on the common sleep process.

Keywords: Movement detection · Respiration rate · Sleep study
FSR sensor

1 Introduction

The importance of sleep in human life should never be undervalued. Our daily routine can be positively or negatively influenced by the quality of sleep. Studies show that the recommended sleep duration varies based on personal body attributes and even more – on the age group. However, the amount of sleep alone does not guarantee the good

quality of a rest [1]. The identification of the sleep stages and their durations, which present sleep cycles, is also necessary for the accurate evaluation of the quality of sleep.

Sleep can be categorized into stages, which can be ascertained by processing the various electrophysiological signals during sleep. The examples of these signals would be Electroencephalography (EEG), Electromyography (EMG) and Electrooculography (EOG). The brain and muscle activity as well as the eye movements can be captured separately using these measurements [2]. Following the method determined by Rechtschaffen and Kales (R-K) [3] the obtained signals are being analysed to identify, which one of the six sleep stages the body is in.

Rapid Eye Movement (REM) and Non-Rapid Eye Movement (NREM) are the two main categories of sleep stages. About 25% of the sleep typically occurs in the REM stage, while the remaining 75% occurs in the NREM [4]. REM, also known as the dream stage, is the stage where skeletal muscle activity is actively inhibited. The intention of this inhibition is preventing the physical manifestation of activities or movements being executed in the dream. The eye muscles during this phase are engaged in rapid movements under the lids, which causes the name of this sleep stage [5]. NREM can be split in four stages of sleep. The first NREM stage can be described as the transition between being awake and sleep. In other words, it is the stage, where the process of falling asleep occurs. A person in this stage of sleep is still a bit in aware of his or her surrounding and can easily be awakened by sounds or other outside influences. This phase usually lasts about 5–10 min [6]. When the second NREM stage is attained, the subject is really sleeping, what means a transition from falling asleep to real sleep is completed. There are some characteristics typical for this stage: a person becomes less conscious of the surrounding, breathing and heart rate become more regular and the body temperature drops. Healthy people spend approximately 50% of their sleep in this stage [4]. NREM 1 and NREM 2 conjunct are also known as light sleep. The combination of NREM 3 and NREM 4 is also called deep sleep. Starting in NREM 3 in the brain the delta waves or extremely slow waves appear to be switching with some faster waves. In this stage, core body temperature decreases even more than in NREM 2, heart rate slows down and the blood pressure drops down. During NREM 4 stage only delta waves are produced by the brain. When a person wakes up during these stages (NREM 3–4), the first feelings can be groggy and disoriented [7]. These stages form a sleep cycle that for healthy adults continues 90–120 min and then repeats again [8].

Sleep is definitely not a waste of time. Research has demonstrated that the brain remains active during the process of sleep. Very important role in brain process such as memory consolidation and brain detoxification is played by sleep. During the person is sleeping, the brainstem, hippocampus, thalamus and cortex help to consolidate different kinds of memorial [9]. Moreover, during sleep, the glymphatic system detoxifies the brain from toxins that were built up while consuming energy during the day [10]. Understanding the process of sleep and analysing the individual human sleep patterns, can help to intervene and improve the quality of the sleep and consequential the life quality. The lack of sleep can result in many health issues such as migraine, insomnia, hypersomnia and in worst case scenario even death.

Nowadays sleep can be analysed processing the electrophysiological signals in sleep laboratories. Unfortunately, the sleep environment in sleep laboratories is very different from home and feeling being sleeping like at home cannot be reached. This makes it therefore even more difficult to obtain results equal to home-sleeping with high accuracy. In addition, continuous monitoring can be easier executed in home environment. Another important factor to be taken into account is cost of sleep analysis. The process of sleep study in sleep laboratory requires large financial and human resources. Respiration and heart rate, but also body movements are important in determining sleep behavior [11]. Furthermore, monitoring body movement in addition to breathing during sleep can support the detection of apnoea and myoclonic [12]. The aim of this work is to develop an efficient system for collecting data about movement and respiration of a person while sleeping and without any additional physical contact to a person's body such as wearable sensors.

2 State of the Art

Sleep study and sleep stage classification as its part has different methods applied for measurement and further data processing. The most widely used method for measuring sleep patterns is Polysomnography (PSG) [13]. PSG includes: Electroencephalography (EEG), Electrocardiography (ECG), Electrooculography (EOG), Electromyogram (EMG). Some other measurements could also be executed as a part of a PSG session, like e.g. video recording.

There are also other methods used in sleep monitoring apart from PSG. One of researches focused on the current state of the art on this field is [14]. Not only the comparison of presented on the market commercial devices with their properties and application fields was executed, but also the short overview of current researches on the topic of health monitoring systems using sensors on beds or cushion was presented in this paper.

Some methods are based on the use of motion sensors, which measure the body movement – actigraphy [15]. Actigraphy can be worn on the wrist or the hip to record sleep and wake behavior [16]. Actigraphy signals can be analysed with much more details, if they are recorded with a high sampling rate [17]. Then it is possible to derive even respiration from actigraphy recording. In most cases an actigraphy-method, pushes data to a computer after some period of time or even offline. Some other proposing studies work with wearable devices and read time data [18]. The continuous wearing of the devices could provoke the feeling of being permanently in an uncomfortable environment, resulting that these methods have limitations in being used long-term. Combinations of methods are also presented in research. An example is the combination of actigraphy and respiration data [19]. The mentioned sleep monitoring methods cannot provide all data in comparison to PSG as described by R-K Method, but they could provide enough information to diagnose sleeping disorders and classify the sleep stages.

Identifying of some ‘non-conventional’ sleep stages, such as the pre-wake stage, are also a topic of research [20]. This is important for quantifying drowsiness before and after sleep as well as during the sleep phase. Here, the patient suddenly wakes

up. The continuous analysis of everyday behavior is very helpful in such cases for detecting symptoms that might indicate sleep disease syndromes or other health issues.

However, using the established method PSG can be very uncomfortable for long-term observation of the patient, and it is costly and difficult to setup for non-experts. This is one of reasons of a new research trend using non-invasive sleep study systems. Some of examples of sensors used are: video motion system, radar sensors [21], optical fiber sensors, piezoelectric sensors, load cells and pressure sensors (e.g. resistive) [22]. Piezoresistive, capacitive, piezoelectric or optoelectronic principles are applied to detect pressure. Another type of pressure sensors is the Force Sensing Resistor (FSR) [23] that was used in several researches on sleep analysis. Comparing to the piezoresistive sensors that cannot hold up the value under long-term use, FSR keeps the output value stable. It consists of polymer thick film and interdigitating electrodes; FSR has a resistance value that changes according to the applied force. The typical range of applied force recognized by FSR is from 0.01 kg to 10 kg. The deviation of the measured values is not more than $\pm 2\%$ in this range. The durability test showed that even under high temperature such as 170 °C and a force test applying 5.44 kg over $\sim 1.5 \text{ cm}^2$, the value of the resistance remains in the tolerance range of FSR sensor. In the most of researches, FSR sensor was used over the mattress.

These FSR sensors can recognize sleep postures. One of researches [24] has not only classified in total eight positions, but also has proven a precision of up to 97%. Here, the project goal was to analyse the posture of bedbound patients for the avoiding of ulcers. Because of the high requirements of the system, this research is for categorizing postures. The algorithms for continually detecting bed posture has been developed with the low-cost pressure mat, consisting of pressure sensors.

Not only movements during the sleep, but also changes in respiration and heart rate are important for the indirect classification of sleep stages. [25] demonstrates that pressure sensors can also be used to analyse sleep states; there are used to observe the respiration and body movements. For the respiration signal validation, such parameter as Respiratory Rate (RR) was defined that results in a parameter the Respiration per Minute (RPM) and the number of apnoea, occurred during the measurement period. This number is used for calculation of the apnea-hypopnea index. Furthermore, RR can also be used in calculating of sleep depth and sleep cycles [26]. The dynamics of RR in terms of statistical properties can be exploited in addition [27]. Using the analysis of short-term and long-term correlation as defined by the random walk theory allows the distinction between NREM and REM sleep stages.

The ECG itself, heart rate and heart rate variability can be used to analyse and detect sleep stages [28]. Sleep stages differ in terms of statistical properties of inter-beat intervals of the heart [29]. This had been done successfully and not only sleep stages could be detected but also sleep disorders could be recognized [28]. Some specific sleep disorders are recognized with a very high reliability such as sleep apnoea [30]. For this identification both, heart rate interval analysis and ECG wave morphology such as R-wave amplitude are combined to achieve a high accuracy. Other disorders are certainly more difficult to identify, such as insomnia.

FSR sensors were also used in [31] for measurement of respiration rate and body movements during sleep. 28 commercial FSR sensors are included in the system and a wireless network is built using ZigBee. The real-time communication between FSR and

computer is ensured. The developed for this system software contains seven views: (1) the value of sensors using color, (2–4) the output signals from the FSR, (5–7) the average value of fields 2–4. Moreover, this research shows a good result in detecting body movement and respiration rate.

3 System Architecture

3.1 Evolution of a System Design

The system design of the mesh sensor network described in our approach started with the first version of the system [32] but soon have been revised in both hardware and software architecture. Previous hardware architecture was based on a series of off-the-shelf microcontroller boards which didn't provide many ADC pins but were easy to use and served well as a proof of concept implementation. They were interconnected using an I2C bus with resistor series based address arbitration - addresses were assigned depending on the voltage measured at daisy-chained voltage divider. An Intel Edison board (trademark of Intel) was used as a mesh network endpoint, which served as a communication relay from the mesh network to external service, which recorded data and served as a visualization and analysis tool. The benefit of such system design was simplicity of implementation of each of the components but system as a whole required many hand-soldered boards and wires to connect sensors to sensor nodes. A large number of sensor nodes and an external computing unit was required to collect and process measurement data. To address these issues, a new system architecture was proposed - off-the-shelf microcontroller boards would be replaced with custom PCBs, which would enable the connection of a larger number of sensors. Hand soldered board functionalities could be transferred to a PCB needed as an external computational unit and local data storage and processing has been enhanced. The implementation of proposed changes resulted in a system that is more cost-efficient, easier to install and manage, and it does not require external devices while retaining the feature of large scale integration via an API.

A sensor mesh design is well-suited in this scenario, in case a system requires a large number of sensors, which can be accessed from a single point. When connecting sensors to nodes, a time and processor power consuming task of AD conversion is delegated from an endpoint to the external processors. Now, a horizontal system scaling can be implemented. Attention has to be paid to how this is done. Too granular scaling requires a high bandwidth for the communication channel and increases latency. In contrast to that, a conglomerated system requires more complex micro-controllers providing more pins and a higher computational power to process data. The best approach is to determine the system granularity according to the hardware structure of the projected solution. In this case, a node is placed between two rows of sensors. It supports up to eight sensors in a row and up to 16 in total. This amount of sensors per row suits the application on a slatted frame of a bed.

The planned coverage requires three nodes, while total bed coverage requires only six nodes. In comparison to the previous system that required 20 boards. The nodes are still communicating with the endpoint via single I2C but address arbitration is now

implemented digitally. Again, this allows easier system scaling and eliminates the need for voltage divider resistor boards. The development platform used to collect measurements now consists of two nodes, each connected to 12 sensors and an endpoint like shown in Fig. 1.

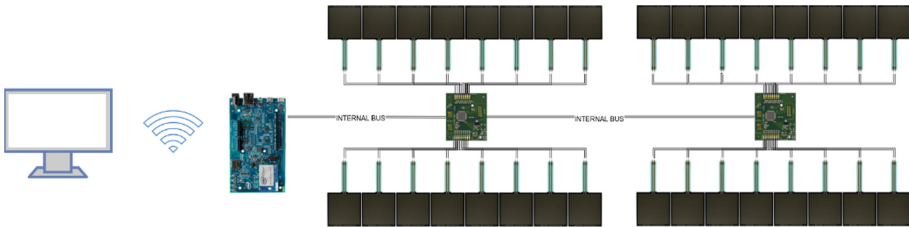


Fig. 1. A new sensor-bed system hardware architecture

3.2 Sensor Node Implementation

A node is implemented as a small and simple PCB. It features an ATMEL SAMD 21 microcontroller based on 32-bit ARM architecture. The benefit of this exact microcontroller is the large number of 16-bit resolution AD pins and the compatibility with a lot of widely-used frameworks and tools. The node PCB uses resistors required for measuring pressure, internal bus connectors for easy interconnection to other boards, a low noise voltage regulator, programming connector, reset button and micro-USB connector. The sensor connector consists of two pins - ground and voltage input pin. 8 pin pairs are located on each side and they are directly attached to the sensors. Programming and debug connector has a cortex debug pinout and is used to flash firmware images and debug. There are two system bus connectors so that boards can be easily chained together. Power is supplied over a system bus or a micro-USB connector and can be connected only on one node. Figure 2 shows board implementation.

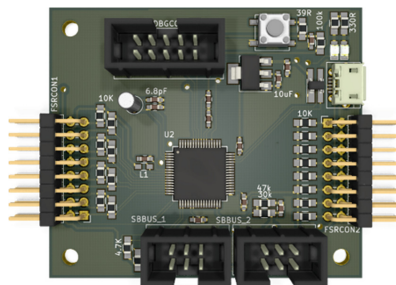


Fig. 2. Node PCB design

The firmware is based on ARM mbed framework and is written in C++. In regular intervals, the node measures voltage value on sensor pins and saves that data in a local dynamic buffer. Sample rate can be increased to improve the time resolution or reduced to save power. When a request arrives via a system bus, the microcontroller processes the request and returns the latest measurements. The node can also change its bus address or work with the endpoint to update addresses of other nodes in case the right request is received. The bus itself is based around I2C communication while two pins are reserved for location sense in and sense out. This architecture allows fully digital address arbitration that starts broadcasting from the endpoint node: all nodes reset their addresses and wait for a high signal on sense input pin. When this happens, the node takes the offered address and responds to the bus. After that, the endpoint instructs that node to rise its sense output pin high so that the next node can catch the address. This arbitration algorithm allows users to place boards in any order and for their sensor mesh network.

3.3 Data Collection

The data collection is done on the endpoint node. Periodically, an endpoint queries the network for the latest data. Received from sensor values are being stored along their timestamps and node locations in a local database. For the beginning, a relational database was chosen as it offers an easy and fault tolerant option for storing data but it turned out to be a bottleneck: the data storage took longer than the sensors' readouts. Nonetheless, it enabled a systematic access to collected data directly on the endpoint. A custom RESTful API was implemented to allow both external services and a visualization interface to read stored data. The API serves JSON messages and it serves information on current system setup and sensor values; as well as previous values determined by either number of reading and/or timestamp.

The implemented data visualization utilizes the aforementioned API to display a live histogram of current sensor values, historic values for sensors on each node and the heat map of the system. The real-time histogram helps users to detect sensor issues during operation. Past values are available for each sensor, while the heath map shows the 'birds-eye' view of the bed at given point in time.

4 Results

Finally, it was important to prove the system architecture for movement and breathing detection. Therefore, a test-system was setup and several experiments have been executed for its evaluation.

24 sensors were connected to the system under the chest area of the mattress. In a first step, experiments were executed with three test persons in different age groups. Two male and one female persons have participated in study. In total, about two hours were spent in bed by each; simulating sleep in different positions have been recorded.

As sensors are placed like a mesh with about 5 cm distance to each other, the signals have differences in amplitudes and they have recorded movements of different parts of body with varying intensity. For simple and clear visualisation at Fig. 3, signals from only three sensors are presented. The signal from each sensor is represented with different colour and this visualizes the change of its pressure value over the time.

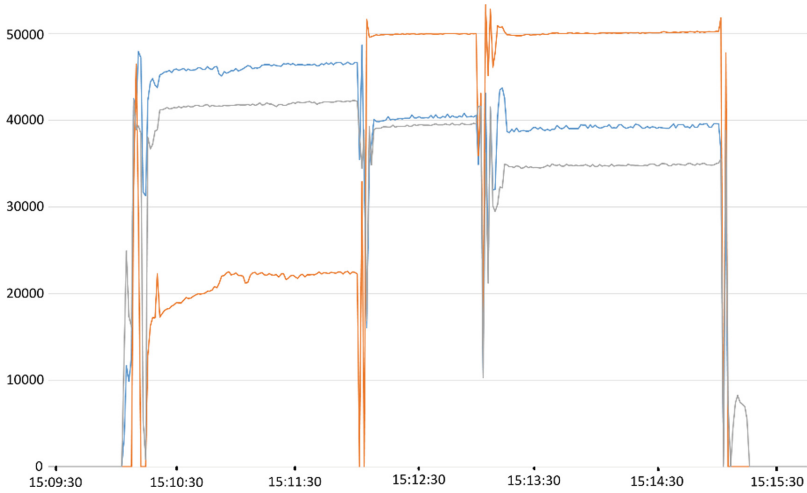


Fig. 3. Signals from 3 sensors (Color figure online)

The distance between sensors represented by the blue and orange line is the maximal one, used in experiment. These two sensors were placed on opposite vertices of a rectangular sensor mesh. From this figure, it is possible to draw the conclusion that in spite of distances between sensors, all of them can record the signal with a sufficiently high sensitivity for recognizing significant movements.

Figure 4 presents data from just one sensor, placed as the top left element of the sensor mesh. It recorded 7 min of an experiment with first sitting on the bed, then lying on it, changing the position and finally getting out of the bed.

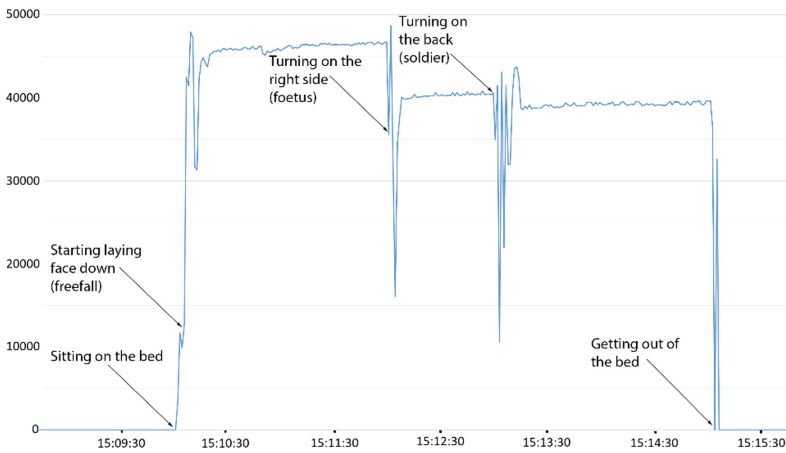


Fig. 4. Test sleep-simulating flow

All movements are clearly recognizable and also the periodical signal can be slightly seen. Figure 5 shows the zoomed-in representation of this periodical signal. It is necessary to mention, that the frequency of signal recording was equal to 1 Hz. This is the main reason, why the periodical signal seems to be not very clear and why it is not possible to recognize exact peaks.

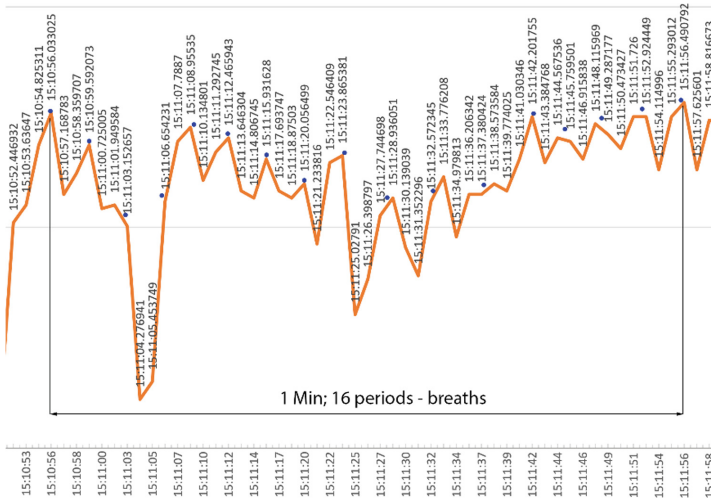


Fig. 5. Zoomed in graph with breathing peaks

As a reference device for evaluating the respiratory rate recognition we used a Zephyr BioHarness - chest belt sensor [33]. Participants have used this device during the test and it shows that the device and the sensors are synchronized with inaccuracy <0.5 s. For respiration rate recognition applied in a sleep study, this is accurate enough, because the data processing will be done for 30 s intervals. Figure 6 shows, the Zephyr BioHarness as a reference, zooming on the breathing signal from for the same period, as in Fig. 5.

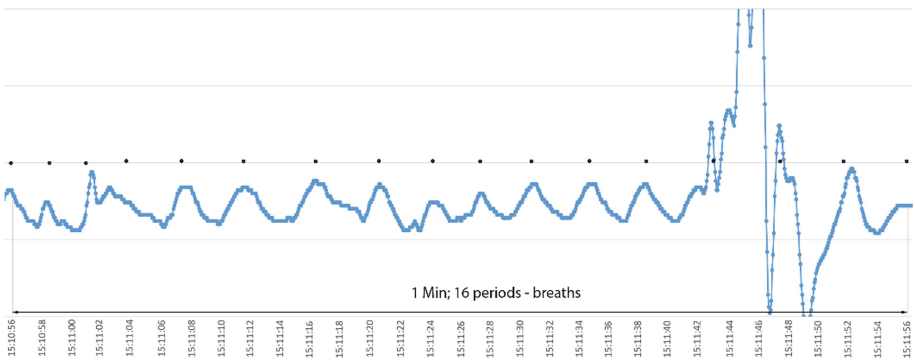


Fig. 6. Reference zoomed in breathing graph from Zephyr BioHarness

The Zephyr BioHarness system uses a different type of sensors and signal processing methods, and therefore, the visualisations of signals shows some differences. The placement of sensors and distance to the measured object – human chest is playing an important role for the recorded and represented signals. Nevertheless, on both graphs, from the sensor bed system and from the chest belt device, the same amount of 16 periods between peaks can be recognized. Of course movement/speaking/etc. should be taken into account to recognize and classify some rare sleep disorders like parasomnias with sleep walking, talking during sleep, or sleep terrors. Therefore, it is not always possible to clearly detect all peaks on a respiration graph, but even with the inaccuracy of 1–2 breaths per minute it could be acceptable for most cases of sleep studies; as breathing rate variability and not respiration rate itself is more important for the conclusions about quality of sleep or its disorders.

5 Conclusion and Future Work

The proposed low-cost system enables measurement of different biovital parameters relevant in sleep phase analysis. The system architecture is composed of a grid of pressure sensors integrated in a common bed. It does not provoke any form of discomfort during sleep since the sensors are placed below the mattress and there no sensors are attached to the body. Body movements and respiration signals could be obtained with help of the sensors. As a consequence, the system seems suitable for sleep analysis providing data related to respiration rate and body movements.

Evaluating the system work with the reference device has confirmed that even with a frequency of 1 Hz, a breathing rate can be recognized. As next step, a long-time experiment with over the night monitoring will be executed and results will be evaluated. In parallel, work on higher scan frequencies have been already started. This may open the possibility to improve the results and enable heart rate recognition. An intermediate step will be the connection to a sleep stage classification algorithm [2] for performing an experiment at a sleep laboratory, where the results can be evaluated in cooperation with sleep medicine experts. The algorithms will include an analysis of high resolution actigraphy recordings in order to derive respiration as shown in previous studies [17]. The algorithms will make use of the dynamics behavior of respiratory rate [27] and of heart rate dynamics [29]. Combining different features from respiratory rate, heart rate and ECG will improve the analysis further to predict sleep stages and sleep disorders [28].

References

1. Hirshkowitz, M., Whiton, K., Albert, S.M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Katz, E.S., Kheirandish-Gozal, L., Neubauer, D.N., O'Donnell, A.E., Oha-yon, M., Peever, J., Rawding, R., Sachdeva, R.C., Setters, B., Vitiello, M.V., Ware, J.C., Adams Hillard, P.J.: National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health* **24**(09), 40–43 (2014)

2. Klein, A., Velicu, O.R., Seepold, R.: Sleep stages classification using vital signals recordings. In: *Intelligent Solutions in Embedded Systems*, pp. 47–50, INSPEC: 15655487 (2015)
3. Rechtschaffen, A., Kales, A.: A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. U.S. Department of Health, Education, and Welfare, pp. 1–12 (1968)
4. Sleep Foundation: What Happens When You Sleep? <https://sleepfoundation.org/how-sleep-works/what-happens-when-you-sleep>. Accessed 04 Dec 2017
5. Sleep Foundation: How Sleep Affects Brain Function. HealthiNation (2012). <https://sleepfoundation.org/video-library>. Accessed 03 Dec 2017
6. VeryWell: The Four Stages of Sleep (NREM and REM Sleep Cycles). <https://www.verywell.com/the-four-stages-of-sleep-2795920>. Accessed 28 Nov 2017
7. National Institute of Neurological Disorders and Stroke: Brain Basics: Understanding Sleep. http://www.ninds.nih.gov/disorders/brain_basics/understanding_sleep.htm. Accessed 05 Dec 2017
8. Sleep Sync: What normally happens during a typical sleep cycle? <http://sleepsync.com/>. Accessed 01 Dec 2017
9. Born, J., Wilhelm, I.: System consolidation of memory during sleep. *Psychol. Res.* **76**, 192–203 (2012). <https://doi.org/10.1007/s00426-011-0335-6>. ISSN 1430-2772
10. National Institutes of Health: How Sleep Clears the Brain. <http://www.nih.gov/news-events/nih-research-matters/how-sleep-clears-brain>. Accessed 02 Dec 2017
11. Muzet, A.: Dynamics of body movements in normal sleep. In: *Sleep*, pp. 232–234 (1988)
12. Nishida, Y., Hori, T., Sato, T., Hirai, S.: The surrounding sensor approach - application to sleep apnea syndrome diagnosis based on image processing. In: *IEEE SMC 1999 Conference Proceedings, 1999 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 6, pp. 382–388 (1999). ISBN 0-7803-5731-0
13. Blood, M.L., Sack, R.L., Percy, D.C.: A comparison of sleep detection by wrist actigraphy, behavioural response, and polysomnography. *Sleep* **20**(6), 388–395 (1997)
14. Orcioni, S., Conti, M., Gaiduk, M., Seepold, R., Martínez Madrid, N.: A review of health monitoring systems using sensors on beds or cushion. In: *Proceedings of International Workshop “Smart Future Living Bodensee”, Konstanz, tbp* (2017)
15. Hedner, J., Pillar, G., Pittman, S.D., Zou, D., Grote, L., White, D.P.: A novel adaptive wrist actigraphy algorithm for sleep-wake assessment in sleep apnea patients. *Sleep* **27**(8), 1560–1566 (2004)
16. Zinkhan, M., Berger, K., Hense, S., Nagel, M., Obst, A., Koch, B., Penzel, T., Fietze, I., Ahrens, W., Young, P., Happe, S., Kantelhardt, J.W., Kluttig, A., Schmidt-Pokrzywniak, A., Pillmann, F., Stang, A.: Agreement of different methods for assessing sleep characteristics: a comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Med.* **15**, 1107–1114 (2014)
17. Wohlfahrt, P., Kantelhardt, J.W., Zinkhan, M., Schumann, A.Y., Penzel, T., Fietze, I., Pillmann, F., Stang, A.: Transitions in effective scaling behavior of accelerometric time series across sleep and wake. *EPL* **103**, 68002 (2013)
18. Velicu, O.R., Martínez Madrid, N., Seepold, R.: Experimental sleep phases monitoring. In: *IEEE EMBS International Conference BHI*, pp. 625–628 (2016). ISBN 978-1-5090-2455-1
19. Long, X., Fonseca, P., Foussier, J., Haakma, R., Aarts, R.: Sleep and wake classification with actigraphy and respiratory effort using dynamic warping. *IEEE J. Biomed. Health Inf.* **18**, 1272–1284 (2013). ISSN 2168-2194
20. Pouyan, M.B., Nourani, M., Pompeo, M.: Sleep state classification using pressure sensor mats. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1207–1210 (2015). ISSN 1094-687X

21. O'Hare, E., Flanagan, D., Penzel, T., Garcia, C., Frohberg, D., Heneghan, C.: A comparison of radio-frequency biomotion sensors and actigraphy versus polysomnography for the assessment of sleep in normal subjects. *Sleep Breath* **19**, 91–98 (2015)
22. Wikipedia, the free encyclopaedia: Pressure Sensor. https://en.wikipedia.org/wiki/Pressure_sensor. Accessed 29 Nov 2017
23. Yaniger, S.I.: Force sensing resistors: a review of the technology, pp. 666–668 (1991). <https://doi.org/10.1109/electr.1991.718294>
24. Pouyan, M.B., Ostadabbas, S., Farshbaf, M., Yousefi, R., Nourani, M., Pompeo, M.D.M.: Continuous eight-posture classification for bed-bound patients. In: 6th International Conference on Biomedical Engineering and Informatics, pp. 121–126 (2013). ISSN 1948-2914
25. Pino, E.J., de la Paz, A.D., Aqueveque, P., Chavez, J.A.P., Moran, A.A.: Contact pressure monitoring device for sleep studies In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4160–4163 (2013). Electronic ISBN 978-1-4577-0216-7
26. Burioka, N., Cornelissen, G., Halberg, F., Kaplan, D.T., Suyama, H., Sako, T.: Approximate entropy of human respiratory movement during eye-closed waking and different sleep stages. *CHEST J.* **123**(1), 80–86 (2003). <https://doi.org/10.1378/chest.123.5.1323>
27. Rostig, S., Kantelhardt, J.W., Penzel, T., Cassel, W., Peter, J.H., Vogelmeier, C., Becker, H.F.: Nonrandom variability of respiration during sleep in healthy humans. *Sleep* **28**, 411–417 (2005)
28. Penzel, T., Kantelhardt, J.W., Bartsch, R.P., Riedl, M., Kraemer, J.F., Wessel, N., Garcia, C., Glos, M., Fietze, I., Schöbel, C.: Modulations of heart rate, ECG, and cardio-respiratory coupling observed in polysomnography. *Front. Physiol.* **7**, 460 (2016). <https://doi.org/10.3389/fphys.2016.00460>
29. Penzel, T., Kantelhardt, J.W., Grote, L., Peter, J.H., Bunde, A.: Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans. Biomed. Eng.* **50**, 1143–1151 (2003)
30. Penzel, T., McNames, J., de Chazal, P., Raymond, B., Murray, A., Moody, G.: Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. *Med. Biol. Eng. Comput.* **40**, 402–407 (2002)
31. Lokavee, S., Puntheeranurak, T., Kerdcharoen, T., Watthanwisuth, N., Tuantranont, A.: Sensor pillow and bed sheet system: unconstrained monitoring of respiration rate and posture movements during sleep. In: IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, pp. 1564–1568 (2012). Electronic ISBN 978-1-4673-1714-6
32. Gaiduk, M., Kuhn, I., Seepold, R., Ortega, J.A., Madrid, N.M.: A sensor grid for pressure and movement detection supporting sleep phase analysis. In: Rojas, I., Ortuño, F. (eds.) IWBBIO 2017. LNCS, vol. 10209, pp. 596–607. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56154-7_53
33. Kim, J.-H., Roberge, R., Powell, J.B., Shafer, A.B., Williams, W.J.: Measurement accuracy of heart rate and respiratory rate during graded exercise and sustained exercise in the heat using the Zephyr BioHarness™. *Int. J. Sports Med.* **34**(6), 497–501 (2013). <https://doi.org/10.1055/s-0032-1327661>



Wearable Pneumatic Sensor for Non-invasive Continuous Arterial Blood Pressure Monitoring

Viacheslav Antsiperov^(✉)  and Gennady Mansurov

Kotelnikov Institute of Radio-engineering and Electronics of RAS,
Mokhovaya Str. 11-7, 125009 Moscow, Russia
antsiperov@cplire.ru, gmansurov@mail.ru
<https://www.researchgate.net/profile/V.Antsiperov>

Abstract. The paper discusses a new type of active sensors for continuous non-invasive monitoring of arterial blood pressure based on the local pressure compensation. Practical implementation of this sensor became possible due to the effective use of modern radio-electronic element base, focused on low power consumption, miniaturization of sizes and built-in computing and communication components (microcontrollers). These characteristics are inherent in modern wearable medical devices, so the sensor proposed can be reliably attributed to this category of appliances. The use of miniature *measuring unit*, its arrangement close to the working area and the possibility of processing digitized data in real time directly in the microcontroller of the sensor made it possible to carry out a unique method of pressure compensation at very small (1 mm^2 or less) areas of elastic surfaces, such as the surface tissues of the human body. The technical implementation of the principle of local pressure compensation in the form of a pneumatic sensor, features of blood pressure measurement regimes, the results obtained and some theoretical recommendations are considered in the paper. The problem of stability of measurement modes is discussed in detail. Experimentally discovered causes of stability disturbance are considered and their theoretical analysis is carried out. The conclusion summarizes the results and outlines the ways of further research.

Keywords: Non-invasive arterial blood pressure (ABP) measurement
Local compensation method · Pneumatic sensor
Microcontroller based sensors · Wearable medical devices
Embedded data processing (EDP)

1 Introduction

People in acute health state [1], such as emergency department patients, intensive care units or operating room patients are often identified as unstable

The work is supported by the Russian Foundation for Basic Research (RFBR), grant N 17-07-00294 A.

(“unwell” [2]), because their physiological condition is characterized as being prone to rapid changes and associated abnormal vital signs [3]. In such circumstances significant additional facilities, resources and personnel are involved in the detection and prompt treatment of injuries and diseases [4]. Operative monitoring of critical physiological parameters is of paramount importance here. The consequences of inadequate patient monitoring may entail “delays in diagnosis, treatment or referral, poor assessment and inadequate or inappropriate patient management” [2] and lead to further morbidity, an increase in the length of stay in the intensive care unit or even mortality.

It is worth noting that monitoring of physiological parameters is important not only in acute states. Due to proper monitoring early signs of a critical illness are detected in time, so acute states can be prevented or even avoided [5]. For example, in the case of cardiac arrest [6], symptoms can be detected a few hours before a serious physiological complication and can be avoided or mitigated.

Currently, the problems of acute care, thanks to technical innovations in the field of (micro) electronics and advances in information technology, acquires qualitatively new features. The ability to acquire data from monitors and bedside systems in real time, the use of the data obtained in support and decision-making applications for prediction, identification and diagnosis of diseases can significantly improve the treatment of evolving medical conditions. In this connection, the main emphasis in the implementation of modern approaches to emergency medical care is shifting to the field of organization and proper processing of information flows of data.

In the paper we discuss a wearable pneumatic sensor for non-invasive arterial blood pressure (ABP) monitoring. It is designed for continuous ABP measurement based on the local pressure compensation method. Practical implementation of this method became possible due to the effective use of modern radio-electronic element base, focused on low power consumption, miniaturization of sizes and built-in computing and communication components (microcontrollers). As mentioned above, these characteristics are inherent in modern wearable medical devices, so the proposed sensor can be reliably attributed to this category of appliances.

The structure of the paper is as follows. Section 2 gives a short overview of existing non-invasive ABP monitoring methods. Section 3 introduces a new method of ABP monitoring proposed by the authors – the local compensation pressure measurement method. Section 4 is devoted to the analysis of local compensation method measurement modes. It contains semi-qualitative estimates of modes stability characteristics and gives some theoretical recommendations to overcome the experimentally discovered causes of measurement instability. Finally, the Conclusion briefly summarizes the results and outlines the ways of further research.

2 Non-invasive ABP Monitoring

Arterial blood pressure is one of the critical physiological parameters widely used both in clinical practice and medical research. Normal ABP is essential

for adequate perfusion of tissues, providing oxygen delivery to all the organs. The ability to measure and monitor ABP trends provides important data about patient cardiovascular status and physiologic state, and helps to define the approach to patient treatment.

Technological advances in the field of microelectronics and computer engineering made it possible in the second half of the 20th century to improve the well-known Riva Roccis and Korotkoffs classic method for ABP measuring [7]. The improvements were mainly connected with the intention to maximally exclude the influence of the human factor on measurements.

Thus, it was found that the most convenient estimate of the pulse for automatic procedures of determination of parameters of ABP is the amplitude of pressure oscillations in an occlusion cuff, a fact that was discovered by Marey a hundred years ago. In 1976 the first bedside automatic instrument for measuring ABP (Dinamap825), which employed a modified method of Mareys, was produced. The transition from a plethysmograph to a typical occlusion cuff combining a device for creation of external compression and a pulsation sensor, which is not very accurate, but acceptable for measuring ABP, gave birth to a new oscillometric technique [8].

The methods mentioned assume a single measurement of the ABP parameters. It is often sufficient for general assessment of the state of the cardiovascular system. However, there are many situations that require continuous beat-to-beat ABP monitoring. At the worst, it is implemented invasively directly inside one of the peripheral arteries upon introduction of a catheter into it. However, since invasive methods relate to patient discomfort and there is a risk of complications upon introduction of a catheter, non-invasive methods are preferable for routine measurements.

Jan Penaz was among the first to put forward a method of continuous non-invasive blood pressure monitoring [8]. His volume clamp method described in 1973 was aimed at reducing the risks of arterial cauterization. The Penaz method employed the idea of “vascular unloading”, based on the assumption that in the “unloaded state” the pressure inside the blood vessel is equal to the outside pressure. The basic element of the device proposed by Penaz is a small finger cuff (Fig. 1A) that has an infra-red (IR) light source on one side and a light receiver on the opposite side. The blood volume of the finger is estimated via the absorbance of IR light. The signal obtained by such a plethysmograph is further used in a feedback loop to control the pressure in the cuff. The pressure is controlled in such a way that blood volume in the finger is kept constant in time and equal to the volume, which corresponds to the unloaded vessels state defined during the calibration process. In this case the oscillations of the controlling pressure are approximately equal to pressure in the arteries. Later some formulae were proposed for recalculating pressure from finger vessels to brachial arteries, which made it possible to verify the method with respect to classical procedures.

Another technique, which provides continuous non-invasive blood pressure monitoring is applanation tonometry [8]. Like the Penaz method applanation

tonometry is based on pulse oscillation estimates, but here the principal of arterial unloading is different. In this case the cuff is placed on the wrist, so the sensor is over the radial artery (Fig. 1B). The sensor presses the artery to the radial bone until it is flattened enough but not occluded. At this intermediate position arterial wall tension becomes parallel to the tonometer sensing surface and arterial pressure is then the remaining stress (perpendicular to the surface) measured by the sensor. The pressure needed to flatten but not occlude the artery is known as the “proper hold-down pressure” and is calculated by a complicated algorithm, which includes the preliminary estimate of systolic, diastolic and pulse pressures over a range of “hold-down pressures” [7].



Fig. 1. Modern approaches to continuous ABP monitoring. (A) CNSystem TM CNAP device, employing the Penaz’s volume clamp method, (B) T-Line device from Tensys Medical TM based on arterial tonometry method

The discussed above methods of continuous non-invasive ABP monitoring have both advantages and disadvantages. Table 1 (reproduced from [9]) presents in a brief manner the pros and cons of the known pressure measurement methods, including the methods considered, in the context of the problems discussed.

Table 1. Brief characterization of known non-invasive methods for ABP measurement

Method	Ambulatory compliance criteria			Clinical compliance criteria	
	Occlusive	Supervision	24 h-prone	Periodicity	Accuracy
Auscultation	Yes	Supervised	No	Intermittent	Gold standard
Oscillometric	Yes	Unsupervised	Yes	Intermittent	Good
Volume-clamp	Partly	Unsupervised	Partly	Continuous	Controversial
Tonometric	Partly	Unsupervised	Fairly	Continuous	Poor
Pulse wave velocity	No	Unsupervised	Yes	Beat by beat	Controversial

3 Local Compensation Pressure Measurement Method

As discussed above, the known methods for the continuous non-invasive ABP measurement are basically reduced to controlling the pressure in the cuff or in applicator compressing the artery along with the limb. Thus, the pressure in the artery is determined according to the measured pulsations of corresponding pressure in the cuff or applicator. Most ABP measurement solutions use for this purpose the principle of volumetric compensation, based on the idea of “unloading the walls of blood vessels”. The principal assumes that in the unloaded state (given by some volume) the pressure inside the vessels (arteries) is equal to the pressure outside them. In the case of a cuff, the pressure is regulated in such a way as to maintain the blood volume in the compressible limb (finger, hand) constant in time equal to the volume which, when calibrated, is selected as “unloading” vessels. In other methods, the applicator presses the artery to the radius bone with such a force that it is sufficient to make flat the surface of the contact with the vessel wall (“flattening” the artery), but not to compress the artery up to occlusion. This is the basis for the sphygmography method.

The disadvantage of the methods mentioned is that when unloading the walls of the arteries when the limbs are compressed, blood stasis is inevitable and intermittent measurements must be periodically interrupted. To solve these problems, we developed a new cuff-less method (see [10] and references therein) of non-invasive continuous measurement of blood pressure, that uses the principle of local pressure compensation [11]. Technically, this method is implemented in the form of a pneumatic sensor construction [12].

Practical implementation of the proposed continuous ABP measurement method became possible due to developing the technique of pressure compensation on very small areas of elastic surfaces (1 mm^2 or less), such as surface tissues of the human body. Technical realization of the proposed method of local compensation was carried out in the form of an experimental sample of a pneumatic sensor [12].

The principle of measuring blood pressure by the method of local compensation and the basic scheme of a pneumatic sensor realizing this method of measurements are presented in Fig. 2A, which also shows the flow of air through the sensor measuring unit during the measurement of pressure P_{art} in the artery.

With the help of Fig. 2A it is easy to explain the idea of local compensation. Namely, if at some initial moment the pressure P_{cam} in the chamber of a sensor measuring unit is less than the artery pressure P_{art} , then the skin and tissue, covering the artery and lying just under measuring unit air channel, keep the output of the channel closed. As the air coming from the receiver accumulates in the chamber, its pressure P_{cam} increases up to the value P_{art} . As soon as P_{cam} reaches P_{art} , the channel output opens and the excess of air escapes under the flat surface of the measuring element connected to the skin. This does not allow P_{cam} to exceed P_{art} and equality $P_{cam} = P_{art}$, continues to be valid for the remainder of the time (even for a quasi-statistically changing blood pressure). Note that the equality $P_{cam} = P_{art}$ automatically maintains the “flattened” condition of the artery wall during measurements.

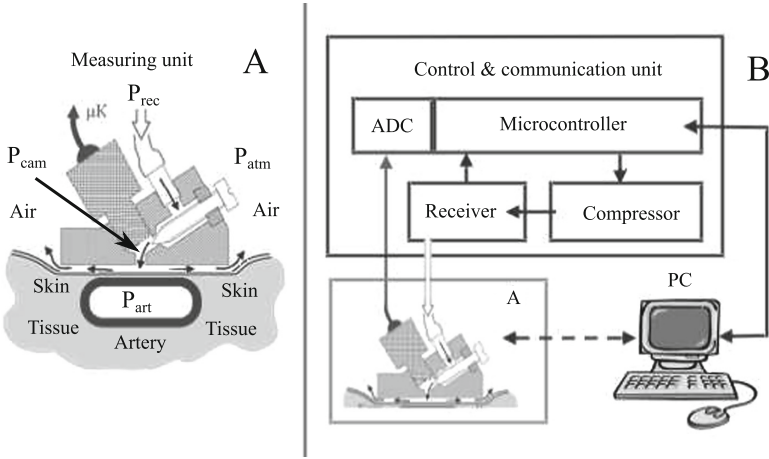


Fig. 2. (A) The principle of local compensation pressure measurement method - if the air steady flows out of the measuring unit chamber, its pressure P_{cam} would exactly coincide with the pressure P_{art} in the artery $P_{cam} = P_{art}$; (B) schematic diagram of a pneumatic sensor with a measuring unit (A) in a sensor feedback loop.

At first glance the operating principle of the pneumatic sensor Fig. 2 is similar to the principle of functioning of the relief valves [13]. However, there is essential difference. If the pressure relief for the valve is a rare event during the normal operation of the system, the pneumatic sensor, on the contrary, operates under conditions of a continuous pressure relief, provided by continuous air flow (and by the blood pressure level in the artery). To provide the desired operating mode, a special design of the sensor measuring unit was developed – it is described in detail by the patent [12]. Below we consider its construction in a very brief manner, only in the volume necessary for understanding the operation of the sensor, and typical measurements that it can provide.

The appearance and the design of the measuring unit of the pneumatic sensor are presented in Fig. 3. The measuring unit Fig. 3A consists of an applicator (10), in the body of which a working chamber (11) is allocated. The working chamber cavity (12) is connected with the pressure sensor (20). The control and communication unit (30) is connected to the output of the sensor (20) via the ADC. The air is supplied to the working chamber (11) from an air pump made in the form of a compressor/receiver pair and connected via an air tube (43) to the high-pressure chamber (44) made in the body of the applicator (10). The chamber (44) is connected to the working chamber (11) via an adjusting choke (45), which in the simplest case is a needle valve.

The applicator (10) has a planar contact surface for installing it in a controlled area of the artery (typically the wrist, see. Fig. 3B). In the center of the contact surface there is a small hole which is an output of air channel connecting the controlled area with the cavity of the working chamber. Through this hole during the measurement, there is a free flow of air.

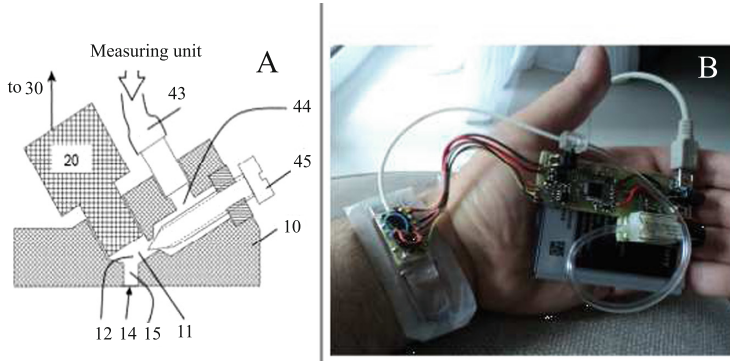


Fig. 3. (A) Design of the sensor measuring unit [12]; (B) Appearance and location of the applicator (10) on the patient's wrist when measuring blood pressure; for illustrative purposes, the control and communication unit (30) is placed in the patient's palm.

Numerous tests have shown that the developed pneumatic sensor is able to monitor the dynamics of blood pressure with a time constant less than one millisecond. This is significantly less than all physiological time constants associated with the dynamics of blood pressure. Typical ABP measurements are presented in Fig. 4.

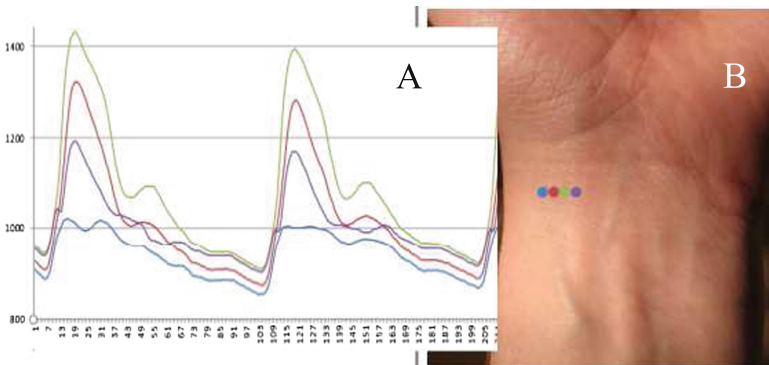


Fig. 4. (A) a typical ABP pulse wave graphics of measured by the pneumatic sensor data, depending on (B) the sensor measuring unit locations in controlled areas (given by different colors) on the patient's wrist.

Pulse waves of blood pressure (representing the dynamics of continuously measured blood pressure) in Fig. 4A were obtained in a specially adjusted mode of the pneumatic sensor operation, determined by several parameters, including the pressure in the receiver P_{rec} , air flow rate, which is also determined by the size of the inlet (adjusting throttle) of the working chamber, the volume of the

working chamber, the size of air channel outlet, controlled area over the artery, etc. The problem requiring such a multi-parameter optimization of the sensor and the measurement mode is that the above mentioned local compensation method ceases to work stably when the quasi-stationary regime of air flow is disturbed (on the time scales typical for ABP dynamics).

Unfortunately, the problem mentioned is not formal, it is quite real. The phenomenon of disruption of stationarity is well known, both in aerodynamics and in the theory of aeroelasticity. In work [14] several mechanisms of possible instability of aerodynamic systems with gas valves are considered. It is shown that the main mechanism for the appearance of non-stationarity is the Hopf bifurcation which usually arises because of the interaction of the valve and the inlet pipe. Bifurcations can lead to different types of aerodynamic turbulence, from isolated jumps in flow regimes to continuous chatter or high-frequency air clatter.

The pneumatic sensor discussed and the presented method of measurement of blood pressure are also subject to disruption of stationarity. In the early experiments, we found empirically that in the air channel from the outlet of the working chamber to the atmosphere formed under the plane surface of the measuring unit, the air flow is stable only at very small cross sections of the channel. This imposes restrictions on the permissible air flow through the measuring unit. Fig. 5 shows a fragment of the experiment, illustrating the instability of airflow with insufficient limitation of air flow rate.

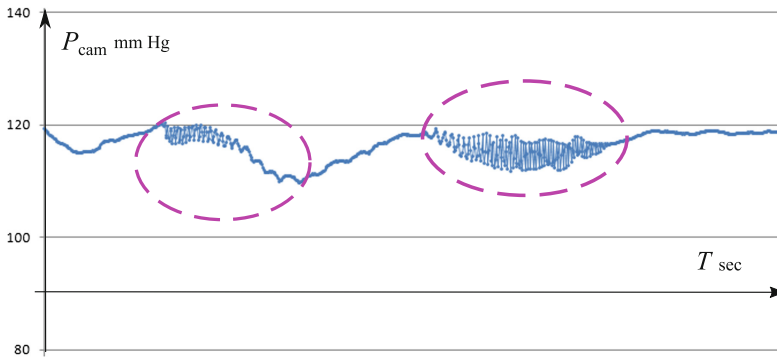


Fig. 5. Disturbance of stationarity regime of air flow, manifested in fragments of high-frequency pulsations, when pressure is measured by pneumatic sensor in an unstable measuring mode.

In view of the above discussed circumstances, it would be desirable to use some theoretical recommendations for the purposeful optimization of the pneumatic sensor and the compensatory method for measuring ABP in addition to the available empirical rules.

Unfortunately, in the literature there are no developed models of radial air flow through flat thin gaps between solid and elastic surfaces. Due to the lack

of a well-developed theory, we present plausible, estimative reasoning, for determining the bottlenecks of the problem under consideration and for working out recommendations for its optimization.

4 Analysis of the BP Measurement Regimens

For quasi-stationary air flow regime, at which $P_{cam} = P_{art}$, let us estimate how parameters of the air flow in the measuring unit depend on the slowly varying blood pressure P_{art} , in particular, to what extent these parameters will change as the blood pressure changes from diastolic P_{dis} to systolic P_{sys} pressure. To begin with, we write the equation of continuity for the air flow:

$$\rho_{in} S_{in} V_{in} = \rho_{atm} S_{out} V_{out} \quad (1)$$

or:

$$\frac{S_{out}}{S_{in}} = \frac{\rho_{in} V_{in}}{\rho_{atm} V_{out}} \quad (2)$$

where ρ_{in} is the density of air in the working chamber, S_{in} is the area of the inlet of the chamber defined by the adjusting throttle (see Fig. 3), V_{in} is the air flow rate generated by the pressure of the receiver P_{res} and coming through the inlet, V_{out} is the free airflow rate in atmosphere, S_{out} – the total area of the gaps of air flow into the atmosphere and, accordingly, ρ_{atm} is the density of air in the atmosphere, corresponding to the atmospheric pressure P_{atm} .

Assuming the air flow is adiabatic (isentropic), we will have $P/\rho^k = const$ with $k = 1.4$ for the air throughout the flow. In this case (2) can be written in the form:

$$\frac{S_{out}}{S_{in}} = \left(\frac{P_{cam}}{P_{atm}} \right)^{1/k} \frac{V_{in}}{V_{out}} \quad (3)$$

We write the Bernoulli equation at the entrance of the chamber:

$$\frac{k}{k-1} \frac{P_{cam}}{\rho_{in}} + \frac{1}{2} V_{in}^2 = \frac{k}{k-1} \frac{P_{rec}}{\rho_{rec}} \quad (4)$$

where ρ_{rec} denotes the air density in the receiver. From (4) follows the expression for the square of the velocity of the incoming stream:

$$V_{in}^2 = \frac{2k}{k-1} \frac{P_{cam}}{\rho_{in}} \left[\left(\frac{P_{rec}}{P_{cam}} \right)^{\frac{k-1}{k}} - 1 \right] \quad (5)$$

Let us write the Bernoulli equations for the leakage of air into the atmosphere:

$$\frac{k}{k-1} \frac{P_{atm}}{\rho_{atm}} + \frac{1}{2} V_{out}^2 = \frac{k}{k-1} \frac{P_{cam}}{\rho_{in}} \quad (6)$$

from which we also express the square of the velocity of the free flow of air into the atmosphere:

$$V_{out}^2 = \frac{2k}{k-1} \frac{P_{cam}}{\rho_{in}} \left[1 - \left(\frac{P_{atm}}{P_{cam}} \right)^{\frac{k-1}{k}} \right] \tag{7}$$

Taking relation (5) to (7), and then the root of the result, we get:

$$\frac{V_{in}}{V_{out}} = \sqrt{\frac{\left(\frac{P_{rec}}{P_{cam}} \right)^\eta - 1}{1 - \left(\frac{P_{atm}}{P_{cam}} \right)^\eta}} \tag{8}$$

where the notation $\eta = (k - 1)/k$ is introduced.

Substituting (8) into (3), we obtain an expression for the ratio S_{out}/S_{in} :

$$\frac{S_{out}}{S_{in}} = \left(\frac{P_{cam}}{P_{atm}} \right)^{1-\eta} \sqrt{\frac{P_{rec}^\eta - P_{cam}^\eta}{P_{cam}^\eta - P_{atm}^\eta}} \tag{9}$$

Introducing the dimensionless variable $\varphi = P_{cam}/P_{atm}$, we finally obtain:

$$\frac{S_{out}}{S_{in}} = \varphi^{1-\eta} \sqrt{\frac{\Pi^\eta - \varphi^\eta}{\varphi^\eta - 1}} = \varphi^{1-\eta} \sqrt{\frac{\Pi^\eta - 1}{\varphi^\eta - 1} - 1} \tag{10}$$

where a dimensionless quantity $\Pi = P_{rec}/P_{atm}$ is introduced.

Since the quasi-stationary mode $P_{cam} = P_{art}$ is assumed, the variable φ can also be considered as a dimensionless blood pressure $\varphi = P_{art}/P_{atm}$. Figure 6 shows, respectively, the form of the dependence (10) of the normalized area of the gaps of air flow into the atmosphere $\sigma = S_{out}/S_{in}$ from the dimensionless blood pressure φ . The threshold depicted at the level $\sigma^c = S_{out}^c/S_{in}$ reflects the experimental fact discussed above about the transition to the unstable regime of air-to-atmosphere flow for large $S_{out} > S_{out}^c$. The intersection of the curve (10) with the threshold σ^c defines the lower limit $P_{art}/P_{atm} > \varphi^c$ of the measuring range of the ABP by the pneumatic sensor.

As for the upper boundary $\varphi < \Pi = P_{rec}/P_{atm}$, or $P_{atm} < P_{rec}$, it is obviously determined by the condition $P_{cam} < P_{rec}$ of air flow from the receiver to the working chamber. Practice has shown that it is better to take P_{rec} with some margin in relation to P_{sys} , since the non linear form of the dependence (10) near Π leads to non linear distortions of the maxima of the pulse wave Fig. 4.

In general, the recommendations for specifying the measurement range P_{art} can be formulated in the following form - choose the interval φ corresponding to the approximately linear fragment of the $\sigma(\varphi)$ (10) dependence, indicated in Fig. 6 as Δ (the rectangle in its boundaries is the current range (P_{dis}, P_{sys})).

In practice, the choice of the recommended range is made, firstly, by selecting P_{rec} to set the upper boundary of the operating pressure range, and secondly by selecting the area S_{in} , which together with P_{rec} determine the air flow rate in the measuring unit and, thus, determines the lower boundary of the range. Namely, if to fix S_{out} at the level of the critical value S_{out}^c , then the lower boundary φ^c will be related to S_{in} by a monotonously increasing dependence:

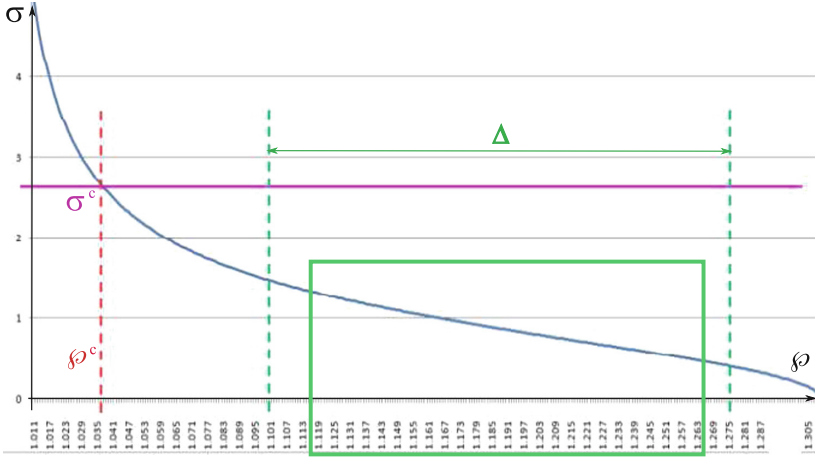


Fig. 6. The dependence (10) of the normalized value of the total atmospheric gaps area σ from the arterial pressure $\varphi = P_{art}/P_{atm}$, normalized to atmospheric. The horizontal line indicates the critical level σ^c associated with the unstable regime of flow.

$$\frac{S_{in}}{S_{out}^c} = \frac{\sqrt{(\varphi^c)^\eta - 1}}{(\varphi^c)^{1-\eta} \sqrt{\Pi^\eta - (\varphi^c)^\eta}} \quad (11)$$

From (11) it follows that in order to reduce the lower boundary of the operating range φ^c , it is necessary to reduce the dimensions of the chamber inlet S_{in} . It has been experimentally established that a decrease in the air flow rate to $Q = 0.1 \text{ cm}^3/\text{s}$ does not lead to a deterioration in the dynamic characteristics of the sensor. At close to the operating mode $P_{rec} = 1050 \text{ mm Hg}$, $P_{cam} = 900 \text{ mm Hg}$ and air density $\rho_{in} \approx 1.5 \cdot 10^{-3} \text{ g/cm}^3$, the air flow velocity at the input of the working chamber V_{in} has, according to (5), a value of the order of 160 m/s . Hence, the cross-sectional area of the inlet choke at an air flow rate of $0.1 \text{ cm}^3/\text{s}$ should be $S_{in} = Q/V_{in} \approx 6.25 \cdot 10^{-6} \text{ cm}^2$, which corresponds to a diameter of 0.03 mm for a round hole. The technology of forming chokes with such characteristics is one of the know-how of the sensor developed.

5 Conclusions

A new method of ABP measurement based on local pressure compensation was proposed. It is shown that the idea of local compensation is quite simple and in part resembles the principle of the relief valve. However, unlike the valve, the pneumatic sensor operates in the regime of constant overpressure relief and to ensure the quasi-stationary nature of this mode of operation it is necessary to ensure conditions for the stable flow of air in the channel formed under the opening of the working chamber and further under the flat surface of the measuring

element. To achieve the stability conditions, it is necessary, as shown in the analytical part of the work, to provide very small dimensions (tens of microns) for certain details of the design of the measuring unit, in particular for the inlet of the working chamber. This uneasy technological problem has been successfully solved by us, the solution is patented [12]. As a result, with the help of an experimental sample of a pneumatic sensor today we have carried out a mass of continuous non-invasive blood pressure measurements, their results are in very good agreement with the results obtained in parallel by known methods.


References

1. Ehrenfeld, J.M., Cannesson, M. (eds.): *Monitoring Technologies in Acute Care Environments*. Springer, New York (2014)
2. Quirke, S., Coombs, M., McEldowney, R.: Suboptimal care of the acutely unwell ward patient: a concept analysis. *J. Adv. Nurs.* **67**(8), 1834–1845 (2011)
3. McGloin, H., Adam, S.K., Singer, M.: Unexpected deaths and referrals to intensive care of patients on general wards. Are some cases potentially avoidable? *J. R. Coll. Physicians Lond.* **33**(3), 255–259 (1999)
4. Sawyer, R.G., Tache, L.C.A.: Common complications in the surgical intensive care unit. *Crit. Care Med.* **38**, S483–S493 (2010)
5. Cuthbertson, B.H., Boroujerdi, M., McKie, L., Aucott, L., Prescott, G.: Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Crit. Care Med.* **35**(2), 402–409 (2007)
6. Franklin, C., Mathew, J.: Developing strategies to prevent in-hospital cardiac arrest: analyzing responses of physicians and nurses in the hours before the event. *Crit. Care Med.* **22**(2), 244–247 (1994)
7. Chung, E., Chen, G., Alexander, B., Cannesson, M.: Non-invasive continuous blood pressure monitoring: a review of current applications. *Front. Med.* **7**(1), 91–101 (2013)
8. Settels, J.J.: Noninvasive arterial pressure monitoring. In: Ehrenfeld, J.M., Cannesson, M. (eds.) *Monitoring Technologies in Acute Care Environments*, pp. 87–107. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-8557-5_12
9. Sola, J.: Continuous non-invasive blood pressure estimation. ETHZ Ph.D. dissertation No. 20093 (2011)
10. Kachuee, M., Kiani, M.M., Mohammadzade, H., Shabany, M.: Cuff-less blood pressure estimation algorithms for continuous healthcare monitoring. *IEEE Trans. Biomed. Eng.* **64**(4), 859–869 (2017)
11. Antsiperov, V.E., Mansurov, G.K.: Continuous non-invasive arterial blood pressure monitor with active sensor architecture. *Int. J. Adv. Life Sci.* **8**(3–4), 289–296 (2016)
12. Antsiperov, V.E., Mansurov, G.K., et al.: Pneumatic sensor for continuous non-invasive blood pressure measurement. RU Patent 2638712, 15 December 2017. Bulletin No. 35. (in Russian)
13. Crosby Pressure Relief Valve Engineering Handbook. Technical Document No. TP-V300 (1997). <http://www.iceweb.com.au/PressRelief/CrosbyPSVEngHandbook.pdf>
14. Hos, C.J., Champneys, A.R., Paul, K., McNeely, M.: Dynamic behavior of direct spring loaded reassure relief valves in gas service: II reduced order modelling. *J. Loss Prev. Process Ind.* **36**, 1–12 (2015)

Healthcare and Diseases



Gene-Gene Interaction Analysis: Correlation, Relative Entropy and Rough Set Theory Based Approach

Sujay Saha¹ , Sukriti Roy¹, Anupam Ghosh²,
and Kashi Nath Dey³

¹ CSE Department, Heritage Institute of Technology, Kolkata, India
sujay.saha@heritageit.edu, sukiroy1357@gmail.com

² CSE Department, Netaji Subhash Engineering College, Kolkata, India
anupam.ghosh@rediffmail.com

³ CSE Department, University of Calcutta, Kolkata, India
kn dey55@gmail.com

Abstract. Logical interaction between every pair of genes in a gene interaction network affects the observable behavior of any organism. This genetic interaction helps us to identify pathways of associated genes for various diseases and also finds the level of interaction between the genes in the network. In this paper, at first we have used three correlation measures, like Pearson, Spearman and Kendall-Tau to find the interaction level in a gene interaction network. Rough set can also be used to find the level of interaction, as well as direction of interaction between every pair of genes. That's why in the second phase of the experiment, entropy measure & Rough set theory are also used to determine the level of interaction between every pair of genes as well as finds the direction of interaction that indicates which gene regulates which other genes. Experiments are done on normal & diseased samples of Colorectal Cancer dataset (GDS4382) separately. At the end we try to find out those interactions responsible for this cancer disease to take place. To validate the experimental results biologically we compare it with interactions given in NCBI database.

Keywords: Gene – gene interaction · Correlation · Entropy · Rough sets
DNA microarray · Colorectal cancer

1 Introduction

A gene, which is considered to be a locus (or region) of DNA, consists of nucleotides and is the molecular unit of heredity [1]. Genes give instructions to create proteins [2]. The transmission of genes to an organism's offspring is the basis of the inheritance. Experiments have proved that most of the characters of the living organisms are controlled by a collaboration of several different genes. This condition where a single character is governed by two or more genes and every gene affect the expressions of other genes involved is known as gene-gene interaction. If this Gene interaction is analyzed properly many valuable information can be mined out regarding the gene expression level variation in different cases of interaction. Gene expression is the process by which information from a gene is used in the synthesis of a functional gene

product [1]. The discovery of gene connectivity/interaction networks from temporal expression data is one of the most pressing problems in computational biology.

Over the decades, researchers are working on the field of genomic data analysis and ranking genes which are both biologically and statistically significant based on a gene microarray experiment.

In this work, we are interested in finding the level of interaction between every pair of genes and also to determine the direction of the interaction, i.e. to identify the regulatory genes in a gene regulatory network. Our aim is also to report how the interactions differ in the diseased dataset from a normal dataset for the same set of genes. We have applied correlation measure, entropy & rough set sequentially to achieve those objectives. Workflow of the whole work is shown below (Fig. 1):

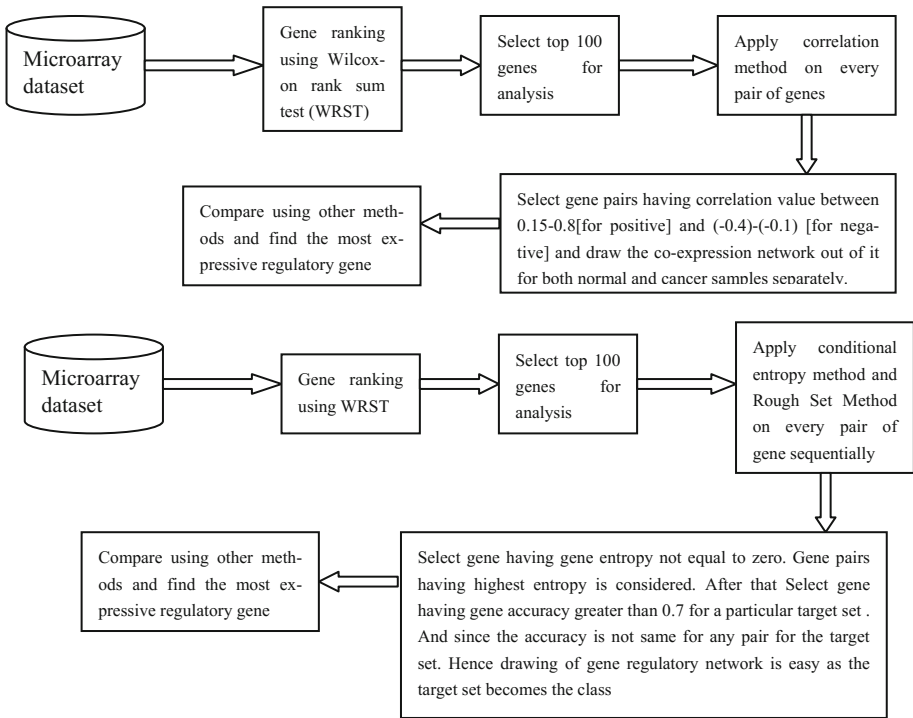


Fig. 1. Workflow of the complete algorithm

2 Literature Survey

Several works have already been proposed to find the level of interactions between genes. Let’s now briefly discuss those methods:

One such work, proposed in [3], used information theoretic measure of synergy, which uses mutual information between the genes. They used the relation between each pair of gene to build the synergy network. High levels of synergy in gene pairs

indicates that two genes shared a pathway and leads to a graphical representation of inferred gene-gene interactions associated with disease, in the form of a “synergy network.” The proposed approach was applied on a set of publicly available prostate cancer gene expression datasets and also validates our results successfully.

A novel algorithm, Signing of Regulatory Networks (SIREN) was proposed in [4] that can infer a regulatory type of interactions for each pair of connected gene of a gene regulatory network (GRN) by computing a similarity score. SIREN score was estimated in four ways in this work, one by B-Spline discretization of expression data, another by calculating co-occurrence scores for each combination of bins of the genes, third one was by rescaling of co-occurrence scores, and the last one was by calculation of expected value of the rescaled co-occurrence probability scores. The proposed approach was applied on three different benchmark gene regulatory networks, including *Escherichia coli*, prostate cancer, and an in-silico constructed network. Experimental results revealed that the new method had approximately 68, 70, and 100 percent accuracy, respectively, for these networks.

Another new method was introduced in [5] for estimating gene group interactions using sparse canonical correlation analysis (SCCA) coupled with repeated random partition and sub-sampling of the gene expression data set. This method infers these types of interactions using appropriate partial correlations between genes. The proposed approach was compared with several existing methods on simulated and real datasets. Experimental results show that the new procedure performs better than those earlier methods in terms of both the statistical measure as well as biological measure.

Another work presented in [6], used differential co-expression analysis of *Saccharomyces Cerevisiae* to build differential co-expression network (DCEN) and identified transcription factors that causes differential expressions under different situations. Result analysis found that differentially co-expressed genes tend to participate in different pathways.

A novel work was presented in [7], where two metrics correlation & entropy were used to find the level of interaction between the genes applied on Gene Interaction networks. Experiments were done on three benchmark cancer datasets Colorectal, Leukaemia and CML. Results show some weighted graphs, where the weights along each edge represents the level of interaction between two genes in a particular network.

A two-stage discovery-confirmatory analysis was proposed in [8] that explored potential gene-gene interactions for hypertension to take place. The first stage was an exhaustive pair wise search performed in 2320 early-onset hypertensive cases with matched normotensive controls from the offspring cohort. In the second stage, identified gene-gene interactions were justified in an independent set of 694 subjects from the original cohort. Experimental results identified four unique gene-gene interactions susceptible to hypertension. Overall this gene-gene interaction analysis helps to identify those genes which can provide more insights into the genetic background of blood pressure regulation.

3 Proposed Methodology

In this work, at first we have used two statistical measures like correlation & relative entropy to find the level of interaction between every pair of genes. Later, a novel meta-heuristic algorithm, named as rough set theory [9] is also used to determine the interaction level. Apart from finding the level of interaction, this work also finds the direction of the interaction, which eventually gives the idea of which gene regulates which other genes. This is done using entropy & rough set principle. We have not found any work till date that uses rough set theory to achieve this. Given the microarray dataset, data is first separated into normal samples and cancer samples. Now from these set of genes, we select top most 100 genes using Wilcoxon Rank Sum test (WRST). We then find interactions between these top most 100 genes. Following sub sections explain those methods one by one.

3.1 Correlation Measure

First analytical approach being used here is correlation measure between a pair of genes. A positive correlation implies that X increases as Y increases and a negative correlation implies that X decreases as Y increases or vice versa. No correlation means that there is no consistent tendency for values on Y to increase or decrease as X increases. In this paper we have done our experiments on three types of correlation measures, like Pearson, Spearman and Kendall-Tau. Let’s explain the use of those correlation measures in the sub-sequent sub sections:

3.1.1 Pearson Correlation Coefficient

Pearson correlation coefficient [10] measures the linear relationship between two datasets. It requires a normally distributed dataset. Let’s now explain this measure & the others also with the help of the following sample dataset:

Table 1. Sample dataset

ZNF85	4.77337	4.55513	4.90434	4.85293	5.06112	5.09473
JPX	5.59942	5.13329	4.8641	6.32837	5.24194	4.49001

The above table consists of 6 cancer samples for two genes, named as ZNF85 & JPX. Pearson correlation coefficient between genes ZNF85 & JPX is calculated as $\text{Corr}(ZNF85, JPX) = -0.31953659412960606$, which will be treated as the level of interaction between these two genes.

3.1.2 Spearman’s Ranked Correlation

Spearman’s correlation [11] is a ranked version of Pearson correlation metric. It does not assume that the datasets are normally distributed. Let’s explain the method we follow in this case using the same sample dataset given in Table 1. It first ranks the dataset and we have the following result:

Table 2. Ranked version of Table 1

ZNF85	2	1	4	3	5	6
JPX	5	3	2	6	4	1

After that we calculate the difference between all the samples of genes ZNF85 & JPX and squares all those differences to get the following result (Table 3):

Table 3. Result after squaring the differences from Table 2

9	4	4	9	1	25
---	---	---	---	---	----

Finally we calculate the sum of all those squared differences and get the interaction level between the given genes as 0.48.

3.1.3 Kendall-Tau Correlation

Kendall tau [12] is used to measure the correspondence between two rankings, where values close to 1 indicate strong agreement and close to -1 for strong disagreement. This measure also ranks the dataset like spearman's correlation and we have the ranked result as Table 2 above. Now it finds the number of concordant and discordant pairs. Subtract those values and divide the difference by $[0.5 * n * (n - 1)]$. For the above set of genes given in Table 1, the level of interaction is 0.2.

The above three correlation measures are applied on the same dataset and correlation is recorded. The percentage of the number of gene pairs having no interactions between them are calculated for both normal and cancer samples.

3.2 Relative Entropy Measure

We then analyzed the same dataset with the relative entropy [13] measure for finding the level of interaction as well as the direction of the interaction in the gene regulatory network. The entropy for every gene pair is not symmetric, so the pair (considering two genes a & b, $a \rightarrow b$ and $b \rightarrow a$) having the higher entropy is considered for the regulatory network.

Let's explain the method with the help of the following example where the p-values of two genes G_1 & G_2 for 3 samples are given as follows:

Table 4. p-values of two genes

G_1	0.41	0.44	0.41
G_2	0.27	0.53	0.14

Relative entropy $H(G_1|G_2)$ is used on the p-values given in Table 4 as follows:

$$H(G_1|G_2) : \sum p(G_1) \log [p(G_1)/q(G_2)] = \sum (0.0744, -0.0356, 0.19) = 0.228$$

Now the same formula is used to find the entropy $H(G_2|G_1)$ as follows:

$$H(G_2|G_1): \sum p(G_2) \log [p(G_2)/q(G_1)] = -0.072$$

Since the value of $H(G_1|G_2)$ is higher than $H(G_2|G_1)$, so we conclude that the interaction will be from $G_2 \rightarrow G_1$ with the level of interaction is 0.228.

3.3 Rough Set Theory

We have proposed an approach using rough set theory to find the level of interaction as well as the direction of the interaction between every pair of genes. The target set is considered as the class or the regulatory gene here. Rough set is calculated for each pair of genes among those top 100 genes selected, considering each gene of the pair as the target set at a time. The pair of gene with its target set having the highest accuracy ensures the regulatory gene, i.e. the target gene is the regulator for the other gene. Let's explain the process with the help of the following categorical dataset $I = (A, U)$ given in Table 5, where A is a non-empty finite set of attributes and U is a non-empty set of finite objects. Here A contains {Gene1, Gene2} and $|U|$ is 10.

Table 5. Categorical dataset

Gene1	4	3	3	2	1	2	3	3	2	6
Gene2	3	3	2	2	2	2	3	2	2	6

Let's first assume that the target set for the class as Gene2 of the above dataset is $\{\{1,2,7\},\{3,4,5,6,8,9\},\{10\}\}$.

Now, for $P = \{Gene1, Gene2\}$, indiscernible objects $IND(P) = \{\{1\}, \{2,7\},\{3,8\}, \{4,6,9\},\{5\},\{10\}\}$.

Next step is to find the lower approximation for $IND(P)$ in the target set. For the above example, lower approximations \underline{P} are as follows (Table 6):

Table 6. Lower approximation

Target set	Lower approximation
{1,2,7}	{1},{2,7}
{3,4,5,6,8,9}	{3,8},{4,6,9},{5}
{10}	{10}

Next step is to determine the upper approximation for $IND(P)$ in the target set. For the above example, upper approximation \overline{P} is as same as the lower approximation. That's why the accuracy of the rough set will be 1. Now the same procedure will repeated by considering target set as $\{\{1\},\{2,3,7,8\},\{4,6,9\},\{5\},\{10\}\}$ where Gene1 is the class.

Now, for $P = \{Gene1, Gene2\}$ this time, indiscernible objects $IND(P)$ will be as same as target set.

Similarly after finding lower approximation like above we have the following (Table 7):

Table 7. Lower approximation where Gene 1 as the target set

Target set	Lower approximation
{1}	{1}
{2,3,7,8}	{2,7},{3,8}
{4,6,9}	{10}
{5}	{5}
{10}	{10}

Similarly upper approximation also gives the same result as above. That's why the accuracy of the rough set will also be 1 in this case too.

Now to determine the direction of interaction, if accuracy is not equal, the object with higher accuracy will be chosen as source. Since in this example, accuracy is same then any of the two objects can be selected as source. Accuracy of the rough set will be the level of interaction between those two objects.

4 Experimental Results

We have applied the proposed approach on Colorectal Cancer dataset (GDS4382) available from NCBI repository [14]. Description of the dataset is given in the following section:

4.1 Description of the Dataset

Colorectal Cancer tumors (GDS4382):

(<http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS4382>)

Organism: Homo sapiens

No. of Genes: 54000 + , No. of Samples: 34 (Normal: 17, Diseased: 17)

Platform: GPL570: [HG-U133_Plus_2] Affymetrix Human Genome

U133 Plus 2.0 Array

Series published on – 20th March, 2012

4.2 Platform Used

We have done our experiments on Python 3.6. Functions of Scipy library of python have been used. Another python library, NetworkX, is also used to draw the gene network.

4.3 Results Using Correlation Measure

In all our experiments, we first separate the normal & diseased samples from the colorectal dataset (GDS4382). What we try to find here is those pair of genes for which there is interaction in normal samples but no interaction in diseased samples. We set a range [0.15–0.8] for positive correlation and [–0.1–(–0.4)] for negative correlation (Table 8).

Table 8. Thirty sample interactions present in normal set, but absent in diseased set

Sl. No.	Gene 1	Gene 2	Correlation value
1	ETV4	ATP6V1E2	0.176471
2	MSANTD3	MTBP	0.279411
3	ATP6V1E2	ZNF674	0.455882
4	SCIN	ATP6V1C2	0.220588
5	ZNF485	TRIB3	0.220588
6	MTBP	ABCC13	0.161765
7	JMJD1C-AS1	SH2D1B	0.426471
8	SLC26A7	EFCAB3	0.308824
9	TIGD1	BC038299	0.235294
10	SH2D1B	BC035642	0.205882
11	BRWD3	TMEM72	0.161765
12	MIER3	SH2D1B	0.470588
13	EFCAB3	MYCBP2	0.205882
14	CDYL2	AU140866	0.25
15	METTL6	DPCD	0.397058
16	AJUBA	PRR26	0.191176
17	DCBLD1	ATP6V1C2	0.191176
18	NKX2-3	CDYL2	0.323529
19	CCDC122	DKFZp451A211	0.184503
20	DPCD	C8orf59	0.411764
21	ATP6V1C2	JMJD1C-AS1	0.191176
22	USB1	DDX42	0.161765
23	SLC4A4	AJ227860	0.382353
24	TTC37	NBEAL1	0.161764
25	BC012528	ZC3H12C	0.161765
26	RBM33	JPX	0.220588
27	TRIM4	CNNM2	0.426470
28	JPX	PHLDB2	0.235294
29	ATP8B3	BC038299	0.264705
30	FBLIM1	ZNF575	0.323529

4.4 Results Using Entropy and Rough Set Measure

Relative entropy and rough set theory are also used here to find the level of interaction between every pair of genes. This time we are also interested in finding the direction of interaction. Experiments are done on cancer & normal datasets separately. For rough set we have set a threshold value of 0.7, i.e. if the accuracy of rough set falls below 0.7 then there will not be an edge between that pair of genes. Following table shows a comparison in direction of interaction between some of the gene pairs in cancer dataset using entropy & rough set.

Table 9. Comparison in measuring level & direction of interaction using entropy & rough set

Number of Gene 1	Identifier of Gene 1	Number of Gene 2	Identifier of Gene 2	Entropy value	Rough set accuracy	Remarks
1671	ETV4	288	JMJD1C-AS1	1.00E-16	0.94117647	
26	MSANTD3		JMJD1C-AS1	2.82E-17	0.64705882	
	MSANTD3	5274	TMCO4	5.05E-17	1	Mismatch
288	JMJD1C-AS1	26	MSANTD3	1.76E-17	0.82352941	
		407	SLC26A7	1.40E-16	0.88235294	
		713	BRWD3	2.92E-17	0.70588235	
		1762	TBC1D16	8.61E-17	0.82352941	
		2645	SPIN3	1.32E-17	0.76470588	
		2957	BC040862	7.08E-17	0.76470588	
		3144	1556657_at	6.13E-17	0.88235294	
		3445	ZNF575	1.63E-16	0.70588235	
713	BRWD3	288	JMJD1C-AS1	1.42E-17	0.94117647	
1096	AJUBA			4.25E-17	1	
32	ATP6VIC2			1.50E-16	0.88235294	
1572	JPX	5274	TMCO4	1.50E-17	1	
1574	MIER3			3.19E-17	0.94117647	Mismatch
1585	ZNF674	288	JMJD1C-AS1	1.27E-17	0.88235294	
2573	TRIB3			1.21E-16	0.82352941	
2667	GON4L	5274	TMCO4	1.10E-17	1	
2716	C22orf42			8.54E-17	0.94117647	Mismatch
2934	BC035063			9.09E-17	0.94117647	Mismatch
3458	ATP13A4			288	JMJD1C-AS1	1.38E-18
4012	STRAP			3.31E-17	0.82352941	Mismatch
4377	ZBTB38			2.27E-16	0.82352941	Mismatch
	ZBTB38	5274	TMCO4	1.84E-16	0.94117647	Mismatch
5189	KIAA0226		TMCO4	4.14E-17	0.94117647	Mismatch
5274	TMCO4	713	BRWD3	9.00E-17	0.94117647	
		1122	NKX2-3	1.26E-16	0.94117647	
		1456	TRIM4	1.60E-16	1	
		1574	MIER3	6.88E-18	0.94117647	

It is shown from Table 9 above that in very few cases direction of interaction mismatches given by relative entropy and rough set theory approach. Following figures show the corresponding network structures. Since it will be cumbersome to show gene identifiers in the figure, we replace it with a gene number. The mapping of Gene identifier and gene number is shown in the above Table 9.

We have also found those common interactions returned by all the three methods. Following Fig. 2 shows two such networks, one each for normal & diseased samples, where the level of interaction is the correlation value and the direction is coming from entropy and rough set measures.

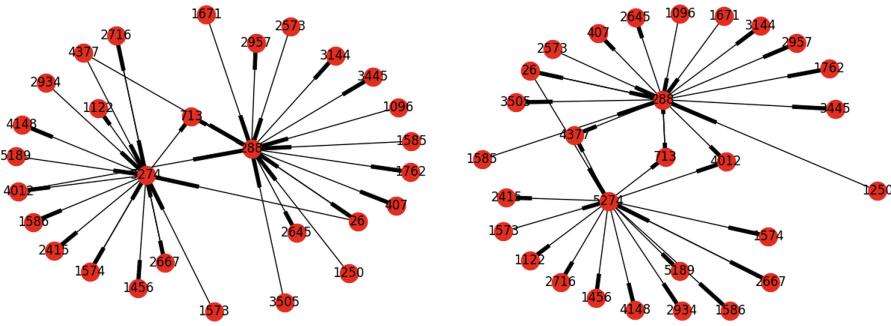


Fig. 2. Corresponding gene regulatory network of Table 9

We have also compared the experimental results with the interactions shown in NCBI standard repository for the colorectal cancer dataset (Fig. 3). If we consider the range of the positive correlation as 0.15 to 0.8 and of negative correlation as -0.1 to -0.4 then we identify all the correlations in the NCBI database within this interval that we find in our experiment. But when we try to find the same match within the top most 500 genes, then we identify 11 such interactions with correlation & entropy within the interval. Following Table 10 shows some of those interactions matched with NCBI database.

Table 10. Some gene interactions matched with NCBI

Number of Gene 1	Identifier of Gene 1	Number of Gene 2	Identifier of Gene 2
1187	CCDC122	2332	CCDC172
1671	ETV4	11669	EP300
26	MSANTD3	11192	LUM
		27904	SUV39H1
		22285	ZDHHC17
32	ATP6V1E2	1250	ATP6V1C2
601	TIGD1	17928	HIST3H3
		22285	ZDHHC17
776	MIER3	11281	HDAC2
4986	ZC3H12C	1083	IRAK2
1017	CDYL2	17990	HIST3H3
4230	BEND5	9946	EIF3D
16154	FOXN2	10159	SKP1
12217	CCNG2	1514	PPP2R5C

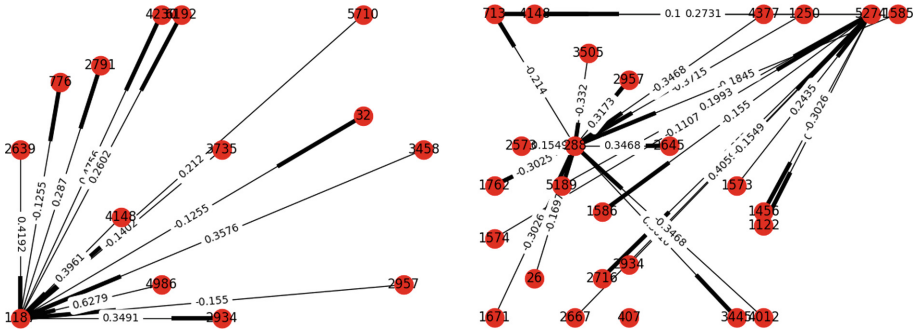


Fig. 3. Common interactions given by three methods for both normal & diseased samples

5 Conclusion and Future Scopes

In this work we have two objectives. One is to find the level of interaction and the other is to find the direction of interaction between every pair of genes present in normal & diseased samples of a benchmark dataset. To achieve these two objectives we have used several methods. We have seen after our experiments that there are some interactions present in normal set, but are missing in diseased sample. What we conclude from this result that may be those missing interactions is responsible for that disease to take place. From finding the direction of interaction we can identify the gene regulation, i.e. which gene or genes are regulated by which other genes. We have compared our experimental results with the interactions provided in NCBI database. Since almost all the interactions we found in the NCBI repository within the interval set by us, it actually validates our experiment biologically.

It is very much clear from the above experiments that identifying these interactions between every pair of genes plays significant roles as far as a specific type of cancer to take place. In our experiments although we find all such interactions matched with NCBI data, but if we consider the top most 100 genes, then only one interaction is matched. That's why to make the process much more accurate we have a plan to apply some strong models like, Rough-Fuzzy or Fuzzy-Rough, Deep Neural Network in future to identify the level of interactions and the direction of interaction as well.

References

1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: Molecular Biology of the Cell, 4th edn. Garland Science, New York (2002)
2. Help Me Understand Genomics– Cells and DNA, Genetics Home Reference (2017)
3. Watkinson, J., Wang, X., Zheng, T., Anastassiou, D.: Identification of gene interactions associated with disease from gene expression data using synergy networks. *BMC Syst. Biol.* **2**, 10 (2008)
4. Khosravi, P., Gazestani, V.H., Pirhaji, L., Law, B., Sadeghi, M., Goliaei, B.: Inferring interaction type in gene regulatory networks using co-expression data. *Algorithms Mol. Biol.* **10**, 23 (2015)

5. Wang, Y.X.R., Jiang, K., Feldman, L.J., Bickel, P.J., Huang, H.: Inferring gene–gene interactions and functional modules using sparse canonical correlation analysis. *Ann. Appl. Stat.* **9**(1), 300–323 (2015)
6. Hsu, C.L., Juan, H.F., Huang, H.C.: Functional analysis and characterization of differential co-expression networks. *Sci. Rep.* **5**, 13295 (2015)
7. Seal, D.B., Saha, S., Chatterjee, M., Mukherjee, P., Mukherjee, A., Mukhopadhyay, B., Mukherjee, S.: Gene — Gene interaction: a clustering, correlation & entropy based approach. In: 2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, pp. 1–6 (2016). <https://doi.org/10.1109/uemcon.2016.7777833>
8. Meng, Y., Groth, S., Quinn, J.R., Bisognano, J., Wu, T.T.: An exploration of Gene-Gene interactions and their effects on hypertension. *Int. J. Genomics* **2017**, 7208318 (2017)
9. Pawlak, Z.: Rough sets. *Int. J. Comput. Inf. Sci.* **11**, 341–356 (1982)
10. Pearson, K.: Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **58**, 240–242 (1895)
11. Spearman, C.: The proof and measurement of association between two things. *Am. J. Psychol.* **15**, 72–101 (1904). <https://doi.org/10.2307/1412159>
12. Kendall, M.: A new measure of rank correlation. *Biometrika* **30**(1–2), 81–89 (1938). <https://doi.org/10.1093/biomet/30.1-2.81>
13. Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951). <https://doi.org/10.1214/aoms/1177729694>. MR 0039968
14. <https://www.ncbi.nlm.nih.gov/>



A Transferable Belief Model Decision Support Tool over Complementary Clinical Conditions

Abderrouf Hadj Henni^{1,2(✉)}, David Pasquier^{1,2}, and Nacim Betrouni³

¹ Cristal UMR CNRS 9189, University of Lille1, 59655 Villeneuve d'Ascq, France
raouf.hadj.henni@gmail.com

² Department of Radiation Oncology, Centre Oscar Lambert, 59000 Lille, France

³ INSERM U1171, University of Lille2, 59045 Lille, France

Abstract. This paper presents an algorithm for decision support over two complementary clinical conditions given a large features data base. The algorithm is mainly divided in two parts, the first one aims at identifying relevant features from a large dimension data base using a heuristic method based on a discriminating power. The second part is a tool based on the Transferable Belief Model (TBM) which combines information extracted from the selected features to provide decision results with probabilities along with a result's consistency measure so that decision could be made carefully. The proposed algorithm is tested on a downloaded feature data base. The TBM based decision support tool showed consistent results w.r.t provided outcomes by combining data from two relevant features identified after using the heuristic feature ranking method.

Keywords: Belief functions · Transferable Belief Model
Decision support · Bio marker identification · Data fusion

1 Introduction

Medical practitioners are often confronted at making decisions between two complementary clinical conditions as in cancer diagnosis or in treatment response prediction. For example, early prediction of treatment response represents a major issue in oncology since it allows the practitioner to adjust the considered treatment w.r.t expected results by analyzing information provided from different bio-markers, however, such practitioner's analysis can be error prone and time consuming.

Biomarkers are usually selected from a data base containing a large number of features extracted from clinical sources such as medical imaging or genomic analysis. In fact, the development of several methods allowing to extract a large

A. Hadj Henni is currently a PhD student at PRISME EA 4229 Univ Orleans, France. However, work in this paper was done during an internship in 2016.

number of features from medical imaging have led to an increased quantitative exploitation of these last. The conversion of medical imaging to mineable data is a process known as Radiomics, and according to [7,8], this process considers that the extracted features can provide precious information concerning cancer detection, diagnostic, and prognosis of treatment response. Genomic analysis can also provide a large number of features which can contain predictive biomarkers [1] that can be useful *e.g.* for personalization of cancer treatment [11] as therapy. However, diagnosis and prognosis rely usually on practitioner's interpretation such as the analysis of the gene expression signature or the visualization of several series of images which could be error prone and time consuming, hence, a tool which can support practitioner's at making decisions is very convenient.

This paper proposes an algorithm which can be used for decision support over two complementary clinical conditions given a large features data base. The proposed algorithm is divided in two parts, the first one is feature selection, which aims at selecting relevant features (*i.e.* biomarkers) from a given data base containing a large number of features using a heuristic method. The second one is data fusion and decision making tool which combines information obtained from the selected features and transforms the combination results into probabilities to help making decisions. Moreover, in contrast to similar studies, our tool quantifies the result's consistency by analyzing the conflict between biomarkers during the fusion process so that practitioners could make a decision carefully.

There exist several feature selection methods for biomarker discovery and different studies have evaluated the performances of some methods as in [3, 4]. However, results obtained in these two studies concerning the most robust method were different. Also, the performances and evaluation criterion vary from one study to another, which proves that there is no selection methods better than others because all depends on the considered study and the related criterion. Note that the aim of this paper is neither to compare existing feature selection methods nor to look for the best one since, we are rather interested only in selecting relevant features from a data base so that to highlight the advantages of the TBM based decision tool when using imperfect and imprecise data. Hence, we propose a heuristic feature ranking method based on discriminating power in order to identify appropriate bio markers. In our case, discrimination means that a given feature should not behave in the same way over different clinical conditions so that a discrimination can be observed on the feature's behavior w.r.t these conditions. Hence, the more a feature is discriminant, the more it is efficient next for the diagnosis or the prognosis task. Relevant features (biomarkers) could be selected then after features ranking as presented further in Sect. 3.

Actually, information obtained from a biomarker is often imperfect and/or imprecise, and to improve decision results, it is rather wise to combine information provided by several biomarkers. Dempster-Shafer theory DST [18] (*also known as theory of evidence*), which is a theory based on belief functions, has been used for data combination in several related clinical applications. For example, DST was recently used for outcome prediction in tumor treatment studies

[13, 14], however, these last concerned features selection rather than providing decision results. In [5, 19], DST was used along with fuzzy sets for medical diagnosis applications, however, conflict between sources (*i.e.* consistency of the fusion process) was not thoroughly analyzed since only mass of the empty set was considered as conflict.

In this paper, we use the Transferable Belief Model (TBM) [16] which is an alternative to DST among the theories based on belief functions and basic differences between them can be found in [2]. The TBM offers several combination operators and allows to transform the combination results into probabilities which makes it more suitable for decision support compared to DST. In contrary to previous related studies, we further analyze the consistency of the combination process using a pair of parameters which is more suitable. Moreover, we define some thresholds, and show how to compute them, so that decision could be made carefully. The rest of this paper is organized as follows: Sect. 2 presents only basic TBM aspects used in this paper, Sect. 3 details the heuristic feature ranking method and the fusion/decision making tool, results and discussion are in Sects. 4 and 5 concludes this paper.

2 Data Fusion and Decision Making Within TBM

This section presents some aspects of the TBM framework which are: knowledge representation, information fusion and decision making.

2.1 Knowledge Representation

- **Modeling information into masses:** Let $\Omega = \{h_1, h_2, \dots, h_n\}$ be a set of hypothesis (*e.g.* clinical conditions), and the power set $2^\Omega = \{\{\phi\}, \{h_1\}, \dots, \{\Omega\}\}$ Information provided by each source (*e.g.* each biomarker) has to be modeled into mass functions $m(H_i)$ which take values in the range $[0, 1]$ and are affected to elements H_i of the power set 2^Ω , with: $\sum_{H_i \in 2^\Omega} m(H_i) = 1$.

Among challenges within belief function theories, including TBM, is the estimation of the mass functions for which the universality does not exist. In our case, information are modeled using an approximate mass distributions method more or less similar to approach used in [17], which depends on the bio-marker’s behavior as explained further in Sect. 3.2.

- **Mass weighting:** Mass functions can be weighted w.r.t reliability of there sources. The basic mass weighting is the one from *Shafer* [18] which consists in multiplying each mass by a corresponding reliability factor $(1 - \alpha)$ where α is a discounting rate. Let m^α be the weighted mass, hence:

$$\begin{cases} m^\alpha(H_i) = (1 - \alpha) * m(H_i). \text{ For all } H_i \text{ except } \Omega \\ m^\alpha(\Omega) = (1 - \alpha) * m(\Omega) + \alpha. \end{cases} \quad (1)$$

2.2 Data Fusion

- **Combination operators:** Basic combination operators within the TBM framework are the conjunctive combination, Dempster’s operator which is a normalized version of the conjunctive operator and the disjunctive operator. Since we are interested in clinical applications involving complementary conditions such as (*tumor/healthy, recurrent/non recurrent, ..etc.*), hence, the frame of discernment $\Omega = \{condition1, condition2\}$ contains absolutely the condition that describes the real status. This case is known as the closed world for which we use the Dempster’s operator:

$$m_{1,2}^\alpha(D) = m_1^\alpha \oplus m_2^\alpha(D) = \frac{1}{1 - k} \sum_{A \cap B = D \neq \phi} m_1^\alpha(A)m_2^\alpha(B) \quad (2)$$

$$With : K = \sum_{A \cap B = \phi} m_1^\alpha(A)m_2^\alpha(B) = m_{\oplus}(\phi) \quad (3)$$

Where: $A, B, D \subseteq 2^\Omega$. m_1^α and m_2^α are weighted masses obtained from bio marker 1 and bio marker 2 respectively. $m_{\oplus}(\phi)$ is the mass of the empty set. In [10] authors showed that when the conflict (mass of the empty set) is not too high, the Dempster’s operator offers good performances w.r.t to other operators. However, when the conflict is too high, the Dempster’s operator led to counter intuitive results, which prompts to consider conflict between sources when combining data using the Dempster’s combination operator.

- **Conflict between sources:** In DST, the mass of the empty set is considered as a quantitative representation of the conflict between sources. However *Liu* [15] has shown, through two examples, that the mass of the empty set alone can not be considered as a quantitative representation of the conflict between two sources, except for particular cases. Thereby, many studies were done to quantify the conflict by measuring distances between masses as in [12] or between pignistic probabilities as in [20]. However, the use of distances alone, depending on situations, can not be an ideal choice because of the three following constraints. The first constraint is that the maximal distance is not necessarily achieved in case of total conflict. The second is that the use of distance for more than two sources is tricky, and finally distance are sensible to the size of the frame of discernment. Nevertheless, in our case the frame of discernment is not important since we are interested in clinical applications involving only two focal elements ($\overline{Condition}, \overline{Condition}$). Concerning the second constraint, we consider that we have only few biomarkers to combine, as it is the case for example in the prognosis of treatment response in prostate cancer using MRI imaging, where the first biomarker can be extracted from DWI series [9], and the second biomarker can be obtained from DCE series [6]. Finally, to deal with the first constraint, we represent the conflict by a pair of paramaters, which are the mass of the empty set and the distance between betting comittement, as defined by *Liu* [15] which is detailed further in Sect. 3.2.

2.3 Pignistic Transformation

Since a decision could be easily supported from probabilities, the TBM framework allows the transformation of the combined masses into probabilities through the pignistic probability $BetP$. A decision is supported then by choosing the singleton corresponding to the highest pignistic probability:

$$BetP(C_i) = \frac{1}{1 - m_{\oplus}(\phi)} \sum_{C_i \in A} \frac{m_i(A)}{|A|} \tag{4}$$

Where: $A \subseteq \Omega$ and $|A|$ is the cardinal of A . C_i is a singleton (*i.e.* $|C_i| = 1$) since this transformation is applied only on singletons.

3 Methods

3.1 Heuristic Features Ranking

This step consists in ranking features w.r.t there discriminating power in order to identify bio-markers presenting high discriminating powers.

Discriminating power estimation for each feature: We denote the discriminating power by $Disc(j)$, and it can be estimated as follow:

$$Disc(j) = \frac{Diff_{Med}(j)}{Range(j)} \tag{5}$$

$Diff_{Med}(j)$: Is the difference between population’s median of each condition of the j^{th} feature. $Range(j)$: Is the range of values of the j^{th} feature.

Note: The estimation of $Disc(j)$ concerns rather features which presents overlapping ranges of the two conditions *e.g.* Fig. 1(a). In case where the two ranges do not overlap, the corresponding feature can be seen as perfect feature (*e.g.* Fig. 1(b)) and the calculation of $Disc(j)$ is not consistent. Hence, before applying (5) on the data base, we shall identify and extract the perfect features if they exist. The identification of this latter is explained here after.

Perfect features identification: The identification is based on a comparison test as follows:

$$F(j) = \begin{cases} Perfect & \text{if } Dist_{max} > RG_{max} \\ Non\ perfect & \text{Else.} \end{cases} \tag{6}$$

Where: $RG_{max} = \max(condition(i)) - \min(condition(j))$ is the range of values of the $condition(i)$ having the maximum range value (*e.g.* condition 1 in Fig. 1(a)). with $i \in \{1, 2\}$ and $condition1 = \overline{condition2}$.

$Dist_{max} = |\max(condition1) - \max(condition2)|$. The distance between the maximums of the two conditions.

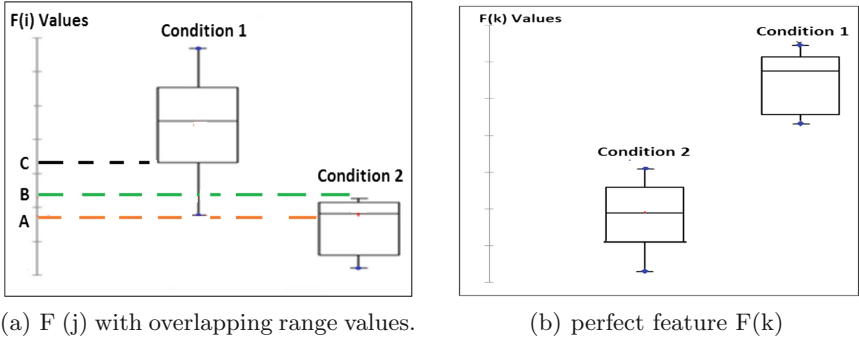


Fig. 1. Boxplots of features

3.2 Bio Marker Data Fusion and Decision Making

Knowledge Modeling: First of all, our Frame Of Discernment (FOD) Ω contains two conditions, $\Omega = \{condition1, condition2\}$, with $condition1 = \overline{condition2}$. This leads to 2^Ω hypothesis: $2^\Omega = \{\phi, condition1, condition2, \Omega\}$. With ϕ represents the empty set, and Ω represents the ignorance. However, since we have a closed world (see Sect. 2.2), then $m(\phi) = 0$. The mass over each condition is modeled as explained in the following example, and the ignorance is then $m(\Omega) = 1 - (m(condition1) + m(condition2))$

For the sake of clarity, let us take the example of a given bio-marker $F(j)$ shown in Fig. 1(a). Information from this latter are then modeled into mass distributions as shown in Fig. 2. A distribution represents a basic belief degree over each condition depending on the value of the considered bio marker $F(j)$.

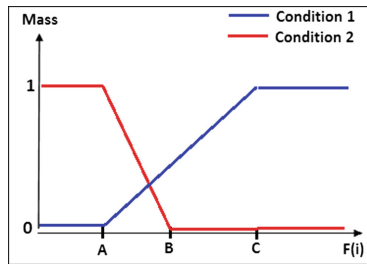


Fig. 2. Mass distribution of a given biomarker $F(j)$

The mass functions over each condition for the considered example can be modeled relying on Fig. 2 as follow :

$F(j) < A$: $m(condition1) = 0$ and $m(condition2) = 1$. Hence $m(\Omega) = 0$

$F(j) > C$: $m(condition2) = 0$ and $m(condition1) = 1$. Hence $m(\Omega) = 0$

$A < F(j) < B$: The mass of both conditions are modeled by a line Eq. (7). Of course, the slope of the line depends on the distribution of the bio marker values over each condition.

$$m(\text{condition}(i)) = A_{\text{Condition}(i)} * F(j) + B_{\text{Condition}(i)}. \quad \text{With } i \in \{1, 2\} \quad (7)$$

Where: $F(j)$ is the current value of the bio-marker $F(j)$, and $A_{\text{Condition}(i)}$ (resp. $B_{\text{Condition}(i)}$) is the line slope (resp. line constant).

$$m(\Omega) = 1 - (m(\text{condition1}) + m(\text{condition2})) \quad (8)$$

$B < F(j) < C$: $m(\text{condition2}) = 0$. For condition1 (resp Ω) as in (7) (resp (8)).

Note: Note that bio marker selection and the knowledge modeling of the second part are performed only once when the complete data base is provided.

Mass Weighting: In our case, more a bio marker is discriminant, more it is reliable. Based on this reasoning, we use the discriminating power as a weighting factor $Disc(j) = (1 - \alpha)$ where α is the discounting rate as shown in Eq. (1).

Biomarkers Data Combination: Combination of the two weighted mass functions obtained from biomarker1 and biomarker2 is performed using the Dempster’s operator defined previously in Eq.(2) where D is in our case condition1 , condition2 or Ω .

Conflict Between Biomarkers: As evoked in 2.2, the conflict between two bio markers is represented by a pair of parameters as defined by Liu [15]:

$$\langle m_{\oplus}(\phi), DifBetP_{m_2}^{m_1} \rangle \quad (9)$$

With: $DifBetP_{m_2}^{m_1} = \max(|BetP_{m_1}(C_i) - BetP_{m_2}(C_i)|)$. C_i could be condition1 or condition2, and $BetP$ represents the pignistic probability defined previously in Sect. 2.3. Liu has discussed the impact of the conflict on the Dempster’s operator result depending on the values of the parameters of the conflict pair w.r.t some thresholds. Nevertheless, no mathematical demonstration was provided to compute them. In this paper, a mathematical demonstration for thresholds computation is provided further in the Appendix. These thresholds will be useful afterwards to discuss the results of the decision support algorithm.

Decision Support and Result’s Consistency Measure: The probability over each condition is provided using the pignistic transformation as follows:

$$P(\text{condition}(i)) = \frac{1}{1 - m_{\oplus}(\phi)} \left(\frac{m_{1,2}^{\alpha}(\text{condition}(i))}{1} + \frac{m_{1,2}^{\alpha}(\Omega)}{2} \right) \quad (10)$$

With $i \in \{1, 2\}$. A supported condition *SuppCond* can then be defined by choosing the condition (i) associated to the highest pignistic probability:

$$SuppCond = \operatorname{argmax}\{P(\text{condition}(1)), P(\text{condition}(2))\} \quad (11)$$

We evoked previously that decision results will be discussed depending on the consistency of the fusion process. This consistency is related to some thresholds defined for parameter values V_i of the pair representing the conflict. In this paper, decision results can be discussed as follows:

$0 < V_i < (0.3 = \varepsilon_{Small})$: Support results could be taken with confidence.

$0.3 < V_i < (0.6 = \varepsilon_{Medium})$: Support results should be taken cautiously.

$0.6 < V_i < 1$: Support results should not be taken.

Where: V_i is the value of parameters $m_{\oplus}(\phi)$ and $DifBet_{m_2}^{m_1}$ separately. The computation of the thresholds is detailed further in the *Appendix*.

4 Results and Discussion

To test our algorithm, we have used a downloaded data base (<http://mldata.org/repository/data/viewslug/colon-cancer/>) which contains almost 2000 features obtained from gene analysis of 62 patients. However, for some patients several data were missing, and for our experiment we kept only 41 patient with almost 400 features just for test. The reduced data base was divided into two bases, a learning data base containing 32 patients, and a validation data base containing 9 patients. *For more details see Excel file in* (https://drive.google.com/file/d/1YPY_kM8CAyzf8zrDx6UezLtp11XFzhHt/view?usp=sharing).

4.1 Features Ranking Results and Biomarkers Selection

Before calculating the discriminating power, we started by looking for perfect features using Eq. (6) over features of the learning data base and we found one perfect feature which is Feature267 (*See line 53 in Excel file*). Hence, this feature is the most discriminant, nevertheless, we did not select it as biomarker since the aim is to show the performances of the algorithm using imperfect biomarkers since in several cases in practice perfect biomarkers can not be found.

We applied thus Eq. (5) for the rest of the features (*see line 55 in Excel file*), and the most discriminant feature was Feature245 with $Disc(F245) = 0.73 \simeq 0.7$. We chose then this latter as first biomarker. The next more discriminating features were F249 and F137 with $Disc(F249) = 0.62$ and F377 with $Disc(F377) = 0.58$. However, those features were more or less correlated to the first biomarker F245. So, in order to diversify biomarkers, we looked for the next more discriminating feature and the less correlated with the first biomarker F245, and it was F241 with $Disc(F241) \simeq 0.4$, we chose then this latter as second biomarker. Note that this latter is less reliable than F245, however, decision results will show that our TBM based tool can deal with this issue. The correlation test was performed using Spearman coefficient since we do not suppose any linear relation, nor distribution law, for the variables (non-parametric test).

To check the consistency of the discriminating power estimation results w.r.t the ground truth, we can rely on boxplots. For example, Fig. 3 shows boxplots of F245, F377, F241 which are discriminant, and the boxplot of F6 which is not discriminant. We can see that the discriminating power values are consistent w.r.t the ground truth, since more the feature is discriminant, more its corresponding $Disc$ is higher. This means that Eq. (5) allows the ranking of features in the right order. However, as we can see in Fig. 3, the real $Disc(F245)$ should be higher than 0.9 (instead of 0.73), same for $Disc(F377)$ and the others. Nonetheless, this latter does not affect the ranking since it is relative to all the rest of the features. Also, the aim of using Eq. (5) is not to calculate the exact discriminating power, but to estimate it so that we can select relevant features.

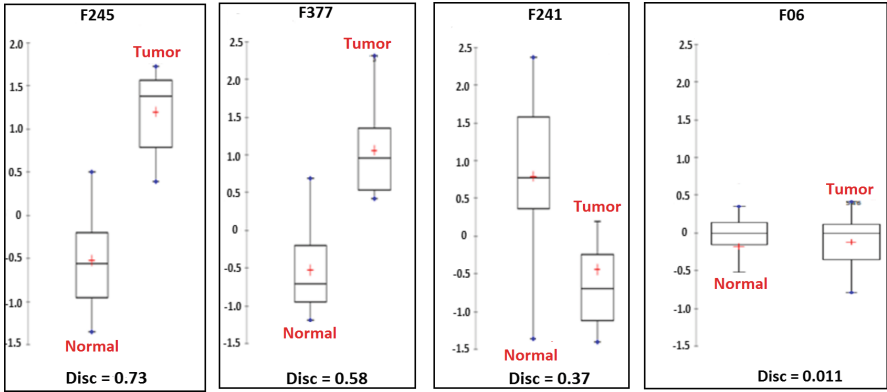


Fig. 3. Box plots for normal and tumor populations for each feature.

4.2 Decision Support Results

Decision results obtained for the validation data base are shown in Fig. 4 and mass functions of F245 and F241 were modeled approximately as shown in Sect. 3.2.

From Fig. 4, we can see that $m_{\oplus}(\phi)$ was null in case of several patients since for those cases biomarkers did not support the *Tumor* condition at all ($m_{F245}(Tumor) = 0$ and $m_{F241}(Tumor) = 0$). However, we can see that their corresponding $DifBet_{m_2}^{m_1}$ was not null, and it was even within medium conflict ε_{Medium} for the first patient *Pat1*, hence, this highlights the interest of using a pair of parameters for consistency quantification instead of $m_{\oplus}(\phi)$ alone.

For the case of patient *Pat4*, we can see that the conflict is medium, nevertheless, decision results are consistent w.r.t outcomes. This consistency is performed since the less reliable biomarker has more or less supported the wrong condition and the more reliable biomarker has supported the right condition. Nevertheless, as we discussed in Sect. 3.2, for medium conflict cases decision results should be taken cautiously in case when the most reliable biomarker will support the wrong condition. Finally, we can see that the decision support results are consistent w.r.t to the provided outcomes.

Patient	F245	F241	Tumoral	Normal	$m_{\oplus}(\emptyset)$	DifBetP
Pat1: Normal	-0,235	0,099	14%	86%	0%	31%
Pat2: Normal	-0,493	1,601	9%	91%	0%	15%
Pat3: Normal	-0,85	1,268	9%	91%	0%	15%
Pat4: Normal	-0,824	-0,621	30%	70%	28%	55%
Pat5: Normal	-0,469	1,837	9%	91%	0%	15%
Pat6: Normal	-1,144	0,664	9%	91%	0%	15%
Pat7: Tumoral	0,417	-0,479	67%	33%	5%	19%
Pat8: Tumoral	1,79	-1,28	91%	9%	0%	15%
Pat9: Tumoral	1,964	-0,377	90%	10%	0%	20%

Fig. 4. Diagnosis results using our decision support algorithm.

5 Conclusion

In this paper, we proposed a heuristic feature ranking approach that helped us selecting biomarkers which were used after by our decision support tool. This latter is based on the Transferable Belief Model and can be used when decision over two clinical events has to be made. Our decision support tool provides decision results with probabilities over each clinical condition and quantifies the result’s consistency by measuring the conflict during the biomarkers data fusion process. In contrary to previous related works, the conflict is rather evaluated through a pair of parameters and its values are then compared to some computed thresholds so that practitioners could make a decision carefully. The algorithm was tested on a downloaded data base and provided consistent decision results w.r.t the provided outcomes despite of data obtained from imperfect biomarkers having different reliability which allows it to be used in practice.

Acknowledgments. First author would like to thank professor *O. COLOT* for his basic belief functions courses given at the university of Lille 1. Work of this paper has been funded by COL (Centre Ocart Lambret, Lille-France) during an internship in 2016.

Appendix: Demonstration for Thresholds Computation

The conflict is divided into three levels, small (trifling) conflict, medium conflict and high conflict. Small conflict is considered when the mass function of the first biomarker over the $condition(i)$ and the mass function of the second biomarker over the complementary condition $\overline{condition(i)}$ are both less than 0.65 (i.e. $m_1(condition(i)), m_2(\overline{condition(i)}) < 0.65$). Medium conflict is considered when those masses are between 0.65 to 0.8, and high conflict when both are higher than 0.8. Note that these values can be chosen differently depending on the application and the expert’s point of view. To estimate the threshold ε for the previous considered values, we proceed as follows:

– **Threshold ε_{Small} for small conflict:**

(1) **Threshold ε_1 for the parameter $m_{\oplus}(\phi)$ of the conflict pair:**

We consider the critical situation between small and medium conflict:

$$\begin{aligned} m_1(\text{condition1}) &= 0.65 & m_1(\text{condition2}) &= a \\ m_2(\text{condition1}) &= b & m_2(\text{condition2}) &= 0.65 \end{aligned}$$

With $\text{condition1} = \overline{\text{condition2}}$ and $a, b \in [0, 0.35]$. It is obvious that the smallest value of $m_{\oplus}(\phi)$ is 0.42, whatever the values of a and b are. Hence the threshold ε_1 for the first parameter $m_{\oplus}(\phi)$ of the conflict pair is $\varepsilon_1 = 0.42$.

(2) **Threshold ε_2 for the second parameter $DifBet_{m_2}^{m_1}$:**

Using the mass functions of the critical case defined in (1) we will have:

$$BetP_{m_1}(\text{condition1}) = 0.65 + \left(\frac{m_1(\Omega)}{2}\right) \quad (12)$$

$$BetP_{m_2}(\text{condition1}) = b + \left(\frac{m_2(\Omega)}{2}\right) \quad (13)$$

Note that:

$$b + m_2(\Omega) = 1 - m_2(\text{condition2}) = 0.35 \quad (14)$$

Hence:

$$BetP_{m_2}(\text{condition1}) = b + \left(\frac{m_2(\Omega)}{2}\right) \leq 0.35. \quad (15)$$

We have also:

$$BetP_{m_1}(\text{condition1}) = 0.65 + \left(\frac{m_1(\Omega)}{2}\right) \geq 0.65 \quad (16)$$

From (16) and (15), we conclude then:

$$DifBet_{m_2}^{m_1}(\text{condition1}) = |BetP_{m_1}(\text{condition1}) - BetP_{m_2}(\text{condition1})| \geq 0.3 \quad (17)$$

Since we have only two singletons (*i.e.* condition1 and condition2), we do not have to look for the maximum of $DifBet_{m_2}^{m_1}$ since it will be the same for condition2. Hence, we obtain:

$$DifBet_{m_2}^{m_1}(\text{condition1}) = DifBet_{m_2}^{m_1}(\text{condition2}) \geq 0.3 \quad (18)$$

From (18) we can see that the threshold ε_2 for the second parameter is 0.3.

(3) **Common threshold ε for both parameters:**

For the sake of ease and precaution, we choose a common threshold ε for both parameters of the conflict pair by taking the smallest value between ε_1 and ε_2 as follows:

$$\varepsilon_{Small} = \text{argmin}\{\varepsilon_1, \varepsilon_2\} = 0.3 \quad (19)$$

– **Threshold ε_{Medium} for Medium conflict:**

Following the same reasoning, we obtain the threshold $\varepsilon_{Medium} = 0.6$.

References

1. Chen, X., Sun, X., Hoshida, Y.: Survival analysis tools in genomics research. *Hum. Genomics* **8**(1), 21 (2014)
2. Delmotte, F., Smets, P.: Target identification based on the transferable belief model interpretation of dempster-shafer model. *IEEE Trans. Syst. Man Cybern. - Part A: Syst. Hum.* **34**(4), 457–471 (2004)
3. Christin, C., et al.: A critical assessment of feature selection methods for biomarker discovery in clinical proteomics. *Mol. Cell. Proteomics* **12**(1), 263–276 (2012)
4. Parmar, C., et al.: Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **5**, 13087 (2015)
5. Wang, J., et al.: A novel method to use fuzzy soft sets in decision making based on ambiguity measure and dempster shafer theory of evidence: an application in medical diagnosis. *Artif. Intell. Med.* **69**(Suppl. C), 1–11 (2016)
6. Roe, K., et al.: Early prediction of response to radiotherapy and androgen-deprivation therapy in prostate cancer by repeated functional MRI: a preclinical study. *Radiat. Oncol.* **6**(1), 65 (2011)
7. Robert, J., et al.: Radiomics: Images are more than pictures, they are data. *Radiology* **278**(2), 563–577 (2016)
8. Kumar, V., et al.: Radiomics: the process and the challenges. *Magn. Reson. Imaging* **30**(9), 1234–1248 (2012)
9. Foltz, W.D., et al.: Changes in apparent diffusion coefficient and t2 relaxation during radiotherapy for prostate cancer. *J. Magn. Reson. Imaging* **37**(4), 909–916 (2012)
10. Liu, Z.G., Cheng, Y.M., Pan, Q., Dezert, J.: Sequential adaptive combination of unreliable sources of evidence (2015)
11. Garman, K.S., Nevins, J.R., Potti, A.: Genomic strategies for personalized cancer therapy. *Hum. Mol. Genet.* **16**(R2), R226–R232 (2007)
12. Jousselme, A.-L., Grenier, D., Boss, L.: A new distance between two bodies of evidence. *Inf. Fusion* **2**(2), 91–101 (2001)
13. Lian, C., Ruan, S., Denoeux, T., Vera, P.: Outcome prediction in tumour therapy based on dempster-shafer theory. In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 63–66, April 2015
14. Lian, C., Ruan, S., Denœux, T., Li, H., Vera, P.: Robust cancer treatment outcome prediction dealing with small-sized and imbalanced data from FDG-PET images. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 61–69. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_8
15. Liu, W.: Analyzing the degree of conflict among belief functions. *Artif. Intell.* **170**(11), 909–924 (2006)
16. Smets, P., Kennes, R.: The transferable belief model. *Artif. Intell.* **66**, 191–234 (1994)
17. Ramasso, E., Panagiotakis, C., Pellerin, D., Rombaut, M.: Human action recognition in videos based on the transferable belief model. *Pattern Anal. Appl.* **11**(1), 1–19 (2007)
18. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, Princeton (1976)
19. Straszcka, E.: Combining uncertainty and imprecision in models of medical diagnosis. *Inf. Sci.* **176**(20), 3026–3059 (2006)
20. Tessem, B.: Approximations for efficient computation in the theory of evidence. *Artif. Intell.* **61**(2), 315–329 (1993)



An Online Viewer of FHR Signal for Research, E-Learning and Tele-Medicine

Samuel Boudet^{1,2(✉)}, Agathe Houzé de l'Aulnoit^{1,3,4}, Antonio Pinti⁵,
Romain Demailly^{1,2,3}, Michael Genin⁴, Regis Beuscart⁴,
Jessica Schiro⁶, Laurent Peyrodie^{1,7},
and Denis Houzé de l'Aulnoit^{1,2,3}

¹ Biomedical Signal Processing Unit (UTSB), Lille, France
samuel.boudet@univ-catholille.fr

² Faculté de Médecine et Maïeutique, Lille, France

³ Groupement Hospitalier de l'Institut Catholique de Lille (GHICL),
Lille, France

⁴ CERIM, EA 2694, Lille, France

⁵ I3MTO, EA 4708, Orléans, France

⁶ Univ. Lille, INSERM, CHU Lille CIC 1403, Lille, France

⁷ Hautes Etudes d'Ingénieur (HEI), Lille, France

Abstract. This paper presents a web viewer of Fetal Heart Rate (FHR) signals developed on HTML5/JavaScript. It provides an easy solution to remotely consult an FHR signal on a web browser without installing any specific software and it can be used either on a computer or on mobile device. There are three major applications of this tool. First, it is used to build up our FHR database used for research on signal processing and analysis. Secondly, it is used on our E-learning website e.RCF in order to train midwives and obstetricians to interpret FHR signals. At last, it could be used on telemedicine either on tele-monitoring to remotely check the fetal welfare, or on tele-expertise to enable practitioners to ask for specialists' opinion. This viewer is designed to correspond to practitioners' habits while including tools to ease the interpretation the FHR signal.

Keywords: Fetal heart rate · E_learning · Tele-monitoring

1 Introduction

The information technologies become increasingly present in healthcare. They provide faster and more reliable access to medical information thus improving healthcare, particularly when several practitioners are involved. Thanks to the development of mobile technologies, information is accessible faster and everywhere. However, it requires appropriate tools to browse information particularly for images and signals.

Web applications are often the most universal and the most convenient way to provide this access on both computers and mobile devices. They have particularly the advantage to not require any installation, to be accessible on any device with the same program, and to easily share any modification on real time which greatly ease

collaborative work (e.g. google drive documents). Since several years, some medical applications have been developed on web support, like some medical image viewers [1–3] which allow slices of an MRI to be browsed, to carry out measures and to annotate them. However, the main drawback is that images are heavy and using those applications is often not as fluent as if it was on more native device language (e.g. C/C++). The image processing is still also limited (image segmentation, normalization seems not yet proposed).

Medical signals could also greatly benefit from those technologies and compared to images, signals have the advantage to be lighter and require less resources for both processing and data transfer. Surprisingly, it seems that this practice is not very developed and the only medical signal web viewers we know are for E-learning [4].

The signal which concerns us on this paper is the Fetal Heart Rate (FHR) which is used to evaluate the fetal wellbeing during pregnancy. The FHR is recorded to monitor pathologic pregnancy (diabetes pre-eclampsia, fetal anemia, ...) and systematically during the labor from 4 cm of cervical dilatation. Recording is most often done on maternity but could also be done at patient's home by either a liberal midwife, or within the framework of in-home care.

To measure FHR in beats per minute (bpm), the device is composed of one or two (for twins) Doppler sensor which can be replaced by fetal scalp electrode in case of signal loss. The device also includes a toco-meter sensor which measures the uterine tonus (to locate contraction), and often a sensor to measure maternal heart rate which is used to differentiate the fetal from maternal heart pulse.

To analyze an FHR during labor, practitioners determine principally the following features. First the FHR baseline, which is the average signal outside accidents. A normal baseline is between 110 and 160 bpm. Then, they determine if there are accelerations (increasing of the FHR over 15 bpm and lasting more than 15 s) and decelerations (decreasing of the FHR over 15 bpm and lasting more than 15 s). A normal FHR has accelerations and no decelerations. Deceleration type could indicate fetal suffering from various etiologies. Then, they measure the variability of the signal which is the amplitude of variation during one minute (normal variability between 6 bpm and 25 bpm). A decision of emergency fetal extraction (caesarian or instrumental) can be made if FHR is pathologic.

The tool presented in this paper is a web application which displays FHR signal. It works either in real time or for already recorded signals. It provides access to the usual tools for signal analysis. It is compatible with any computer, tablet or smartphone and with any modern web browser. A quiet similar tool has been developed by Neoventa® for education [4] but it is in fact a wide-image viewer. It does not allow users to realize measures on the signal and since an entire image is heavier than the signal samples the transfer could be slow for long recordings.

Our team works on the FHR analysis and particularly on the automation [5] and the E-learning [6]. We needed an FHR web-viewer for various parts of this project (acquisition, annotation, presentation, exercise creation, ...). We have then carried out a first version of this viewer in Flash. Unfortunately, Flash player being progressively deprecated and being unusable on mobile devices, we have remake the program on HTML5/JavaScript. The flash viewer being used during the 5 last years, this new one benefit from an expertise and numerous additional development already implemented.

After description of functional features of this tool, we will describe three major applications which we have tested or we are setting up: research, E-learning, and tele-medicine.

2 System Description

This program which is conceived for general use should have features for all applications. This paper focuses on the viewer which is a part which have to be included in a complete website and this last will manage the signal database, security access and legislation concerns. The program consists to download FHR files from the server (potentially on real time), to display the various signals (FHR, toco, maternal heart rate) and to provide the tools to analyze and annotate this signal. An example of displaying on a smartphone is given on Fig. 1.

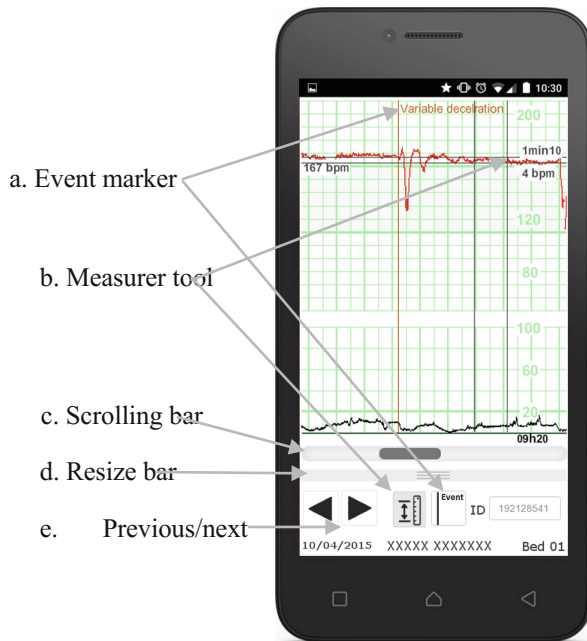


Fig. 1. Overview of the viewer display on an android smartphone

2.1 Conception

The entire program is composed of a JavaScript file and a CSS file. The html webpage has to include those two files as well as some standard library (jQuery, Bootstrap and jDataView) and the viewer can be positioned thanks to an only tag `<fhr-viewer>`. This tag includes some attributes like the FHR filename, and others depending of applications.

FHR files are dynamically downloaded thanks to an AJAX query which enable the viewer to start displaying the signal even if it is not completely downloaded. Since, to our knowledge, there is not real standard FHR file format, we have created our own. The data are binary stored on .fhr files. The first 4 bytes are used to store the UNIX timestamp on unsigned integer. Then for each time sample (at 4 Hz) a package contains first FHR, the second FHR, the maternal heart rate (each multiply by 4 and stored on uint16) then the TOCO value (multiply by 2 and stored on uint8). This file format is defined with the same precision as the data sent by devices. An optional second file, fhrh, stores in text format a piece of information on the subject as well as various event markers.

In some cases, the file has to be downloaded only partially. We have created a streaming service in PHP in which it is possible to specify the beginning and the end of part to download from the server. It is useful for real time monitoring to query only the samples which have not been downloaded yet. It is also useful on E-learning to create progressive clinical case as explained on part IV. We also use this PHP program to check user identification and access limitation before sending the signal thus increasing security.

Once the signal downloaded, an HTML5 canvas frame is generated which displays the grid, signals and event markers. The canvas is generated only for the current displayed time frame to avoid overloading memory. When the user scrolls or when the height of the viewer is changed, the canvas is redrawn. In our tests, on mid-range smartphones the scrolling was fluent so no further optimization seems really required.

The source code is object oriented and includes principally 4 classes: FHRViewer (main class), FHRSIGNAL (a model class which store signal data and manage downloads), FHRGraph (class which display grid, signals, measurer and event markers on the canvas layer) and ScrollBar. This code respects the last W3C recommendations which should reduce compatibilities issues in the next coming years.

The Fig. 2 illustrates the conception of this program for both client and server side.

2.2 Features

The list of all current features of this viewer is given below:

- Works on any computer with any OS (Windows, Linux, MacOS), on any tablet and smartphone (Android, iOS or windows phone). It works on any modern browser (Requirements are chrome v24, internet explorer v9, firefox v21, safari (desktop or iOS) v6.1, android browser 4.4). For older versions on desktop computer (particularly internet explorer 7–9) it is possible to use the previous version in flash.
- Does not require any installation. Once identified, access to recoding can be done thanks to a hyperlink, which greatly ease integration on information systems.
- Display the same grid as paper to fit the user habit (Fig. 1).
- Display one or two fetal heart rate signals, the maternal heart rate and the TOCO.
- Can display, and edit event markers and propose a preselection of standard one. This preselection can be customized depending on the user language and the application (E-learning, clinical...) (Fig. 1a).

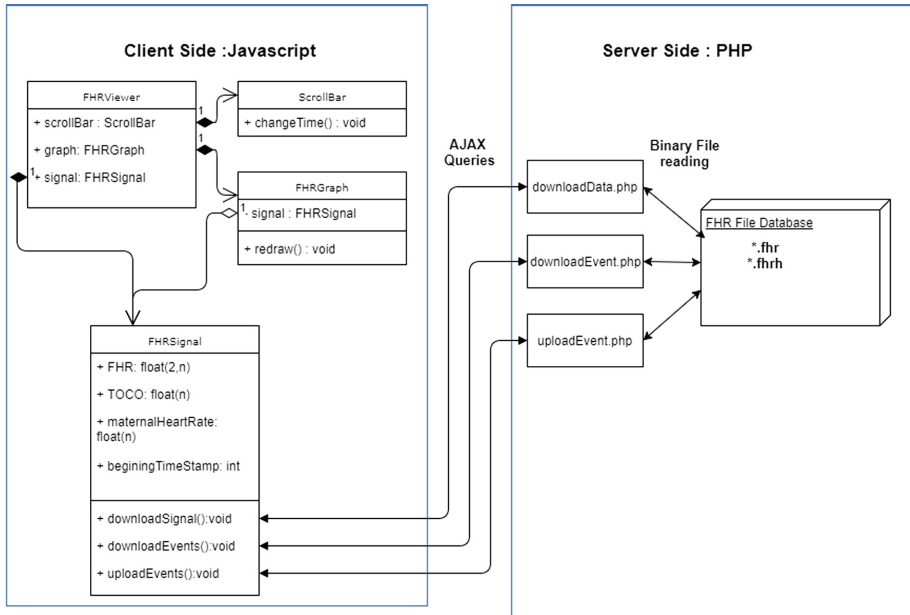


Fig. 2. Conception diagram of the software showing the UML classes and the major interaction

- Horizontal scrolling either by elevator (Fig. 1c) or by an inertial touch move. Next page and previous page button are also available for fast reading. (Fig. 1e).
- Can work on real time for patient being recorded (latency around 1 s).
- Possibility to display several FHR viewers on the same html page (useful for central monitoring).
- Include a tool to easily measure average baseline, acceleration/deceleration duration and amplitude, the time between contraction and deceleration and the FHR long term variability. (Fig. 1b).
- Optimize the viewing scale. Midwives and obstetricians are used to analyse FHR on paper print. We observed that the interpretation of FHR was highly altered (particularly for the estimation of variability) when the signal was not displayed on the same size (1 cm for 20 bpm and for 1 min or 3 min depending on practice). A resize bar has then been included to add possibility to set the height of the graph to fit a paper one while keeping aspect ratio (Fig. 1d). It is also possible to display on 1 cm/min or 3 cm/min depending on preferences.

3 Research Applications

3.1 Data Acquisition

Our research on signal analyses [5] required access to the raw signal data. Signals from the manufacturer central are not always accessible due to security reasons.

Consequently, parallelly to this project, we developed a daemon which generates accessible FHR files from the Ethernet packages circulating between various fetal monitors and the central. To build up a database, those files have to be linked to a patient file and potentially have some annotations. To ensure the systematic and good quality of data-input, viewing the signal in real time is a requirement. Integrated on a webpage, this viewer offers some convenient accessibility possibilities and can be easily customizable depending of needs. Since 2011, the previous version of this viewer enables us to create a database of 8000 recordings. We are also developing a Raspberry Pi® device to transmit on WIFI or on 4G the fetal signal from the cardiocograph to a cloud which will allow us to acquire data from ambulatory recordings.

3.2 Collaborative Research

The most powerful feature of web technologies is to allow multiple users to work on the same files and to instantly share any input to others. Numerous types of annotations could be considered depending on research and it is easily customizable. As example, one of our use was to input the beginning and the end of the second phase of delivery on each recording. Another use we are working on is to realize a comparative study of FHR analysis between several experts.

3.3 Presentation

This tool can be useful to present case studies which are currently often presented by image capture but it is not really convenient for long recordings. With specific authorization, it is even possible to share the file on internet (e.g. additional content of a paper). Researchers on FHR signal processing can also use this tool to present their results (e.g. automated FHR analyses, evaluation of custom indices of variability...).

4 E-Learning Applications

Since 2012, we have settled a French E-learning website called e.RCF (<http://ercf.univ-catholille.fr>) to train midwives and obstetricians at the analysis of FHR signals during labor. There is now 5500 midwives and obstetricians, either active or student, which have tried this website, and 21 francophone midwifery schools used it regularly to train their students. This website includes 150 pages of courses, 100 pre-recorded multiple choice questions, 200 exercises of FHR analysis and 40 progressive clinical cases of FHR. The major part of this website is on free access for training users, and it proposes a school mode to realize examination sessions [6].

- This website used originally the previous flash version of our viewer which has been replaced by this HTML/JavaScript one and thus it is now possible to access this website on mobile devices. There are two types of exercises for which our viewer is perfectly designed:

- FHR analysis: on approximately 40 min of recording, student have to determine the elementary features of the signals (baseline, variability, accelerations and decelerations) and thus deducting objectively the risk of acidosis without knowledge of patient file which tends to bias this analysis. A sample of this exercise can be seen of Fig. 3.

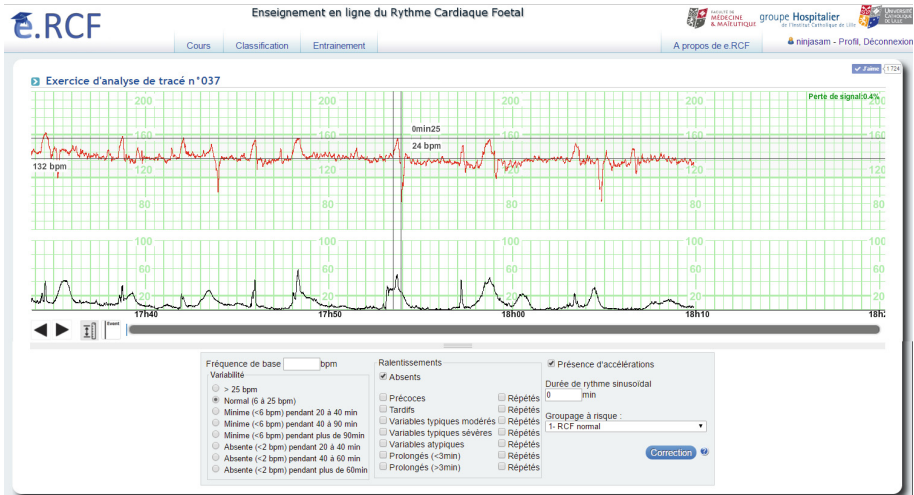


Fig. 3. Web viewer. Overview of an analyse exercise on the French website e.RCF

- Progressive clinical cases: This time, all clinical information is provided. The student has on a first time only the beginning of recording and should combine its FHR analysis with the clinical context to decide the right conduct to adopt. Then he can see the following of the records and decide again of the new conduct taking into account the new elements.

The web viewer offers a very friendly user tool close to the ones used on clinical practice which is generally very appreciate by students. The educational interest of the support has even been evaluated and show a significant improvement of student mark of 2 points/10 for student trained on this website compared to student trained on paper [6].

5 Tele-Medicine Applications

Using this viewer as web application could probably have great benefit on tele-medicine. The French health high authority highly recommend the development of such practice which ease the simultaneous mobilization and increase reliability of treatment process. Unfortunately, our structures do not allow us to get FDA or CE certification to use it on clinical practice. However, we are setting up the tools to realize some tests.

Important requirements of tele-medicine are that any specialist who give a medical opinion has to have access to the entire patient file and should be sure that the quality of image is not degraded. Most of the time, the information systems do not have simple

way to send the signal and consequently users can send data by unreliable solutions like e-mail, telephone and fax. This can create security issues, and reduction of medical data quality which could endanger patients. A better and simpler way is to send a hypertext link (with an authentication system) to the expert which includes both clinical information and the FHR signal. The expert has then the same data and the same tools as if they were on the site. We can cite multiple scenario where this viewer could ease and improve the patient care.

5.1 Tele-Monitoring

A delivery length on average around 8 h and consequently midwives and obstetricians are not close to the patient every time. Particularly, depending of the level of maternal care, the obstetrician on duty can be outside maternity and come only in case of problem. This viewer, being accessible on any web browser and on mobile technology, practitioners can access faster and more regularly to the FHR and thus could intervene faster.

5.2 Tele-Expertise

FHR is generally analyzed by midwives and in case of important anomalies, obstetrician is called for an eventual intervention. It is important that midwives can easily ask opinion of obstetrician and currently the obstetrician have to come in the monitoring room to see the signal. This need is particularly important when recording at patient's home.

This system will also greatly ease the possibility for an obstetrician to ask the opinion of an expert in another center since it is only a hyperlink to share.

5.3 Integration on Hospital Information System

French legislation specifies that FHR signal have to be stored during 15 years. Using device specific software is then problematic since in 15 years we cannot be sure to still have the computer and the software which can open the FHR files. Web technologies are less object to obsolescence since they do not depend of Operating System. Moreover, updates are often simpler since only the server have to be updated and not every client. At last, it is also very easy to attach a FHR signal on the patient file since it is only a hypertext link. With manufacturer software, the only way would be to attach the PDF file of the signal which would be less permissive.

5.4 Sharing the Recording with Patient

Last application, this system could allow the patient to have access to their files. It could be useful for example in case of maternity changes to ease the transfer of a patient file.

6 Conclusion

We have presented a web viewer of FHR signal and presented advantages in different scenarios. FHR signals being light enough, there could be no inconvenient to use such a web viewer instead of device specific software and the accessibility would be greatly improved. We hope that either manufacturer or Health Information System could integrate it in their solutions. Some evaluation of conviviality, medical and comfort benefits and cost interest could be realized to encourage this integration. Our future works would be to include an automated analysis of the FHR on this viewer as well as the displaying of variability measure and the ST Segment Analysis. Moreover, it could be interesting in some cases to have a web tool which make the link between FHR and the partogram. There are probably other medical signals which will benefit from a web viewer and we hope encouraging their development.

Acknowledgment. This project has been partially funded by the “Institut Thématique Santé Publique” and the “Institut Thématique Technologies pour la Santé” within the framework of the call for proposals on connected device.

References

1. MedDream. <http://demo.softneta.com>
2. Haak, D., Page, C.-E., Reinartz, S., Krüger, T., Deserno, T.M.: DICOM for clinical research: PACS-integrated electronic data capture in multi-center trials. *J. Digit. Imaging* **28**(5), 558–566 (2015)
3. Shen, H., Ma, D., Zhao, Y., Sun, H., Sun, S., Ye, R., Huang, L., Lang, B., Sun, Y.: MIAPS: a web-based system for remotely accessing and presenting medical images. *Comput. Methods Progr. Biomed.* **113**(1), 266–283 (2014)
4. Neoventa Academy. <http://www.neoventa.com/education/>
5. de l’Aulnoit, A.H., Boudet, S., Demailly, R., Peyrodie, L., Beuscart, R., de l’Aulnoit, D.H.: Baseline fetal heart rate analysis: eleven automatic methods versus expert consensus. In: 38th Annual International Conference of the IEEE - EMBC, pp. 3576–3581 (2016)
6. Demailly, R., Boudet, S., Houzé, A., Delgranche, A., Houzé, D.: Évaluation d’un programme d’apprentissage en ligne de l’analyse du rythme cardiaque fœtal. *Vocat. Sage-femme* **13**(110), 26–30 (2014)



Modeling Spread of Infectious Diseases at the Arrival Stage of Hajj

Sultanah M. Alshammari^(✉) and Armin R. Mikler

Department of Computer Science and Engineering, University of North Texas,
Denton, TX, USA

sultanahalshammari@my.unt.edu, armin.mikler@unt.edu

Abstract. During the 2009 H1N1 influenza pandemic, there was rising concern about the potential contribution of international travel and global mass gatherings on the dynamic of the virus. The travel patterns after global mass gatherings can cause a rapid spread of infections. Studying the impact of travel patterns, high population density, and social mixing on disease transmission in these events could help public health authorities assess the risk of global epidemics and evaluate various prevention measures. There have been many studies on computational modeling of epidemic spread in various settings, but few of them address global mass gatherings. In this paper, we develop a stochastic susceptible-exposed-infected-recovered agent-based model to predict early stage of a disease epidemic among international participants in the annual Hajj or pilgrimage to Makkah (also called Mecca). The epidemic model is used to explore several scenarios with initial reproduction number R_0 range from 1.3 to 1.7, and various initial proportions of infections range from 0.5% to 1% of total arriving pilgrims. Following an epidemic with one infectious per flight, the model results predict an average of 30% infectious and 20% exposed individuals in Makkah by the end of the arrival period. The proposed model can be used to assess various intervention measures during the arrival of international participants to control potential epidemics in different global mass gatherings.

Keywords: Global mass gatherings · Disease control
Infectious diseases · Epidemic · Outbreak · Agent-based model · Hajj

1 Introduction

Following the identification of the novel influenza A (H1N1) virus in human in Mexico and California in late March 2009, the novel pandemic virus was detected within a few weeks in 20 different countries with the highest volume of passengers arriving from Mexico [14]. The public health officials have a great concern about a further spread of the 2009 pandemic influenza A H1N1 at mass gatherings. Global Mass gatherings (MGs) such as Olympic Games, FIFA World Cup, Hajj, international conferences and musical festivals attract millions of people

from different countries. The travel patterns of the international participants attending these events can contribute to the spreading of infectious diseases and pose public health hazards to the hosting countries and the home countries of the attendees. There have been many studies focusing on the health hazards in global mass gatherings and the need of advanced surveillance systems and international collaboration to control communicable and non-communicable diseases in such massive gatherings [1, 8, 21, 23].

There are several studies on computational modeling of epidemic spread in various settings, but few of them address global mass gatherings. Computational models can be used to have a better understanding of the spread of infectious diseases in these complex settings. Epidemic modeling can provide an efficient tool to study the contribution of global MGs to infectious diseases outbreaks and assist public health officials responded through different control measures [8]. Recent approaches to model the spread of infectious diseases can be classified into two major types, agent-based models [6, 18, 19] and population or metapopulation models [4, 24]. These modeling methods can be utilized in mass gatherings to simulate the mobility patterns of the participating populations and model the spread of epidemics in these events [20, 22]. Ajelli et al. [2] have applied the two modeling methodologies to predict the progression of a baseline pandemic in Italy, using the agent-based model to simulate the Italian population and the metapopulation model to integrate the airline travel data and human mobility patterns. The aim of the their study is to provide a detailed comparison of the limitations, advantages, and assess the achieved results by the two models. The study has demonstrated an agreement in the results obtained by agent-based and metapopulation models in predicting the directions of an epidemic spreading. As agent-based approaches capture the details on individual level, the heterogeneity among the different sub-populations and mobility patterns are integrated using the metapopulation approach. The two methodologies can be combined into a hybrid framework to capture the spread of epidemics on multiple levels at local and global scale [2].

In this study, we are focusing on characterizing the early stage of potential epidemics at global MGs by developing a data-driven computational model for disease transmission at the arrival stage. The global event of Hajj is selected for this study. Every year, more than two million pilgrims arrive from over 186 countries to preform the Hajj at the holy city of Makkah. The influx of such a huge crowd at a confined area for a specific period of time increases the risk of infectious diseases outbreaks. Advanced planning for the Hajj is conducted by the Saudi Ministry of Hajj and Ministry of Health. Health regulations and vaccination requirements are issued and updated prior to every Hajj season to minimize and control the importation of infectious diseases during the arrival of international pilgrims [3, 15]. The event of Hajj provides a challenging setting useful to study and model the impacts of mass gatherings on global spread of infectious diseases. Hajj is very organized religious event with a restricted schedule of rituals at specific holy sites. International pilgrims cannot come to Saudi Arabia to preform Hajj without having been issued a specific Hajj visa and

proof of vaccination for diseases specific for their country of origin. Furthermore, Hajj visas restrict the movement of international pilgrims to the two holy cities of Makkah and Madinah. Hence, it is possible to identify, track, and simulate the movements of international pilgrims during the different stages of the Hajj.

While the actual Hajj rituals can be performed between five to six days, international pilgrims may arrive 30–35 days before the starting date of the Hajj. The long duration of the arrival stage of the Hajj provides sufficient time for a potential epidemic to develop and thrive. The proposed model combines air travel data, population demographic, and data related to the arrival locations and procedures. Stochastic SEIR (susceptible-exposed-infected-recovered) agent-based model is used to simulate the interactions of the international pilgrims and the spread of an epidemic among them. Different scenarios are simulated to assess the impact of the initial size of the epidemic and the starting date of the disease activities among the arriving pilgrims. As this study aims to explore the spread of an epidemic at the arrival stage of Hajj, we did not apply prevention measures or assigning healthy behavior such as wearing a mask to any individual in the population. In the following section we describe the arrival of international pilgrims to Makkah and the related procedures. Section 3 provides a detailed description of the epidemic model and the associated parameters and assumptions. The experimental setup is described in Sect. 4, and obtained results are discussed in Sect. 5. In Sect. 6 we conclude the presented work in this paper.

2 International Pilgrims Arrival

International pilgrims come to Saudi Arabia by air, land, and sea through different entry points. The majority of foreign pilgrims (92–94%) are arriving by air. Air-travel data of five Hajj seasons (2010–2014) are obtained from the Saudi General Authority of Civil Aviation (GACA). The travel data include a daily profile of exclusive Hajj flights arriving to and departing from the two international airports in Jeddah and Madinah. The arrival process starts slowly with a few flights (6–14 flights per day), and increases steadily until it reached its peak (90–150 flights per day) during the last week of the arrival period. The highest increase in the number of incoming flights occurs just days prior to the beginning of the Hajj rituals. Unlike other major events, pilgrims cannot miss the starting day of the Hajj rituals. In addition, the majority of the international pilgrims are arriving from low-income countries, and they cannot afford the cost of accommodation to stay at the holy sites for a long time before the actual rites begin. Therefore, this pattern of flights arrival is not expected to change significantly in future Hajj seasons. Table 1 lists the total number of flights and international pilgrims arriving to Jeddah and Madinah airports of the Hajj season from 2010 to 2014. We also obtained relevant demographic data from focused studies for the same Hajj seasons published by the Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research¹.

¹ www.uqu.edu.sa/en/hajj.

Table 1. Total number of flights and pilgrims arriving to Jeddah and Madinah airports for five Hajj seasons (2010–2014).

	Jeddah		Madinah	
Year	No. of flights	No. of pilgrims	No. of flights	No. of pilgrims
2010	2993	938758	1248	361615
2011	3010	926878	1215	369199
2012	2626	841192	1321	406357
2013	1710	559198	1291	405516
2014	1616	517009	1353	421267

Upon their arrival, the international pilgrims go through several stages including waiting at arrival halls, immigration, customs, baggage claim, and departure area. The average waiting and service times to complete these stages are ranging between 30 to 90 min. However, it may take several hours to complete all arrival procedures in the peak arrival days (20,000–50,000 pilgrims per day) due to the vast number of arriving pilgrims. At the departure area, counters for Hajj missions and agents are located where pilgrims are classified based on nationality into six different establishments using color-coded labeling system. While external pilgrims follow the same procedures in the two airports, those arriving in Jeddah must depart to Makkah once they finished the airport procedures. Foreign pilgrims entering the country through Madinah airport may stay in the holy city of Madinah for several days (up to one month), then travel to Makkah to complete the rituals of Hajj.

During their stay at Makkah or Madinah, international pilgrims will have different accommodations where pilgrims belonging to the same establishment tend to have the same housing arrangements. In both cities, pilgrims move back and forth between their housings, grand mosque, and common areas such as shopping centers and public transportations. On a daily basis, all pilgrims go either to the grand mosque in Makkah or to the Prophet’s mosque in Madinah for the five prayers throughout the day and night at specific times. While pilgrims can be assigned to different locations randomly, all pilgrims who are currently in either Makkah or Madinah will be located in the corresponding grand mosque at five distinct time slots per day. To capture these mixing patterns among arriving pilgrims, several assumptions and parameters are included in our model.

3 Epidemic Spread Modeling

3.1 Social Mixing and Contact Rate

The number of contacts and the mixing patterns are important factors when modeling infectious disease spread [17]. Most of the existing studies explored the contact rates and patterns in confined settings such households, schools, and workplaces. However, there is a lack of information about the mixing patterns

in settings with high population densities such as airport terminals and mass gatherings. Due to the few insights of the contact rate and patterns in these settings, different assumptions of the contact model can be explored. One can assume uniformly weighted contacts and homogeneous mixing patterns for the whole population throughout the different stages of a global event. In reality the contact patterns at global events are unlikely to be homogeneous and the contact rate can vary from one stage to another. However, there does not exist sufficiently detailed data to support other modeling assumptions.

The variations in contact patterns at global MGs arise from population density, demographic and cultural differences, and the event schedule and setting [13, 16]. The characteristics of the individuals can influence their mobility and social mixing. For instance, participants from the same country or who speak the same language have a higher probability of mixing and interacting with each other for longer period of time. At the arrival stage of the Hajj, there can be various social mixing and contact rates. As throughout the different procedures at the airport, pilgrims arriving in the same flight or within the same time frame are often moved together and forced to wait in a close proximity in confined spaces. Also upon their arrival, the international pilgrims are clustered into sub-populations based on their countries. However, these clusters of pilgrims may mix together in various locations during their stay in Makkah or Madinah. The varied number of contacts and interactions between individuals in the arrival stage are generated based on the following assumptions:

1. As only the contacts between an infected individual and a susceptible individual may lead to the transmission of a disease, contacts among non-infectious individuals need not be generated or tracked.
2. A contact model is developed based on the features of the infectious individual and the expected density of the current location in which social mixing take place.
3. A homogeneous mixing within each group of pilgrims arriving from the same country and stay in the same location.
4. We adopt the spatial contact rate model derived in [13] and used by [25] to calculate the contact rate as:

$$C = k\pi p_0(1 - e^{-r^2 \frac{p}{p_0}}) \quad (1)$$

Where $k(0 < k < 1)$ is a constant fraction of contacts among the population, r is the maximum infectious radius of the disease being modeled, p is the average population density of current location, and p_0 is the maximum population density in the crowded situation. p_0 is selected in the range of 7–10 persons per m^2 based on a detailed study tracking the movements of pilgrims at the grand mosque during Hajj season [10] and, population densities extracted from observed data at Hajj [13].

5. In each time step an infectious individual interacts with C_s susceptible individuals, where $C_s = C \frac{S}{N}$, the proportion of contacts from the susceptible individuals S in a population of size N and contact rate C .

- The probability of choosing a susceptible individual for a contact will be quantified by several factors including arrival flight, country of origin, establishment, gender, and age-group.

3.2 Stochastic SEIR Agent-Based Model

Agent based modeling is used to simulate the interactions of international pilgrims during the arrival stage of the Hajj. Each agent represents a pilgrim and maintains a record of their demographic (e.g., gender, country of origin, age), health information (e.g., transmissibility, susceptibility, number of days being exposed or infected), and vaccination status based on the modeled disease(s). Each agent can follow health-related behaviors such as wearing a mask, sanitation, or distancing from symptomatic persons. Other parameters related to the event such as location, stage, arrival date and time, are assigned to every pilgrim. The location parameter tracks the current location of pilgrims where it can be assigned based on the stage or ritual they are performing, or it can be determined based on a likelihood of remaining at the same location or move to a different location(s).

The SEIR compartmental model is used to classify pilgrims into compartments based on their infection status. The movement of individuals between these compartments reflects the propagation of the disease in the population. Figure 1 shows the timeline of an infectious disease in humans. When being in contact with an infectious individual, a susceptible individual will move to the exposed compartment, where the individual is infected but not yet infectious. At the end of the latent period, those individuals will be capable of transmitting the disease, and will be moved to the infectious compartment. In the beginning of the infectious period individuals start as asymptomatic then they show symptoms of the disease. At the end of the infectious period, infectious individuals will recover, gain immunity to the disease, and move to the removed compartment. In our SEIR model we include both asymptomatic and symptomatic infectious compartments and currently individuals in both compartments have the same likelihood to infect other susceptible individuals. Having the two disease states will allow future updates for the epidemic model. For instance, the transmission probability can be quantified and associated with being asymptomatic or symptomatic for some disease.

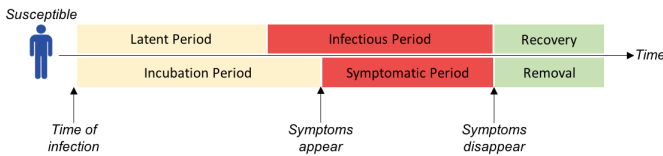


Fig. 1. The natural history of the progression of a disease in an individual over time.

4 Experimental Setting

The disease model in our experiments is parameterized using the epidemiology of influenza H1N1. The incubation period of influenza H1N1 lasts up to 4 days and the duration of the infectious period is 7 days [26]. To calculate the contact rate using Eq. 1, we used the infectious radius of influenza as approximately ≤ 6 feet [11]. The basic reproduction number R_0 for influenza H1N1 is estimated in different studies in the range of 1.2 to 2.3 [5, 12, 26, 27]. The value of R_0 is defined as the average number of newly infected cases generated by one infected individual during the infectious period in a fully susceptible population. R_0 depends on the average contact rate, duration of infectious period, and the probability of transmission (transmissibility).

We modeled the epidemic spreading in different scenarios at the arrival stage. Each scenario is simulated for 100 Monte-Carlo realizations for each set of values of the model parameters. In each realization, the stochastic parameters associated with disease transmission changed. The average of the number of cases across these realizations is calculated. These scenarios are represented by changing the following simulation settings:

1. R_0 , the basic reproduction number.
2. p_t , transmissibility, the probability of infection transmission per a contact between a susceptible and infected individual. In case of following a healthy behavior (e.g., wearing a mask), the transmission probability will be reduced to p_h , ($0 < p_h < p_t$). Transmissibility is determined using duration of the infectious period and the average contact rate to meet the initial value of R_0 [7, 9].
3. λ ($0 < \lambda < 1$), the likelihood of a flight arriving with one or more infectious individuals.
4. I_0 , the proportion of infectious individuals per flight, per day, per sub-population, or per the whole population of pilgrims.
5. t , the starting date of the epidemic.

Epidemics are initiated by seeding the population at the arrival to the main entry points with one or more infected individuals. Those individuals can be selected randomly from the whole population, or from sub-populations defined by arrival date, country of origin, or age-group. Different scenarios are simulated to study the factors affecting the initial size of the epidemic. In these scenarios, the initial size of the epidemic is defined based on a probability that a particular flight will arrive with one or more infectious individuals. The starting day to initiate the epidemic can be the first day of arrival or randomly selected within the arrival period. When a flight arrived with one or more infective, individuals are selected randomly from the passengers of that flight and their disease status will be assigned based on a parameterized probability as exposed, infectious asymptomatic, or infectious symptomatic. The latent or infectious timer is set for that individual within the range of the latent and infectious periods of the disease being modeled.

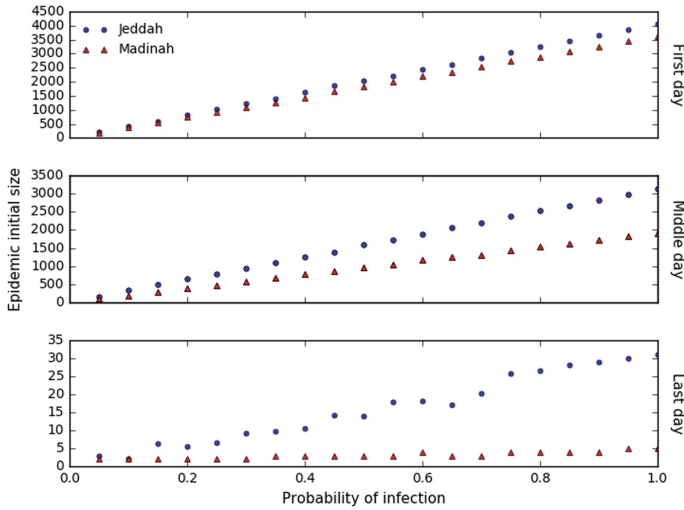


Fig. 2. Estimates of the initial size of the epidemic among the international pilgrims within 35 days at different starting dates with one infectious per flight. In each graph, each point represents the total of infectious individuals over the entire population of international pilgrims at the end of the arrival period.

5 Results and Discussion

5.1 Initial Epidemic

To explore the effect of the likelihood of a flight arrives with I_0 infected individuals on the initial size of the epidemic, we executed several simulations using different values of λ per a simulation. The value of λ starts with 0.05 and increased by 0.05 every simulation. We also, execute these simulations with different starting days, $t =$ first, middle, or last day of the arrival period. Different proportions of initial infectious individuals were explored ($I_0 = 1$ or $k\%$, where $k = 0.5, 1, 1.5,$ and 2). Figure 2 shows the average of each simulation for the total number of infected and exposed individuals over the entire population of arriving pilgrims at the end of the arrival period. The results obtained under these scenarios showed that initial epidemic size is influenced by the value of λ and the starting date. As the value of λ increases, the initial size of the epidemic increases in both arriving populations at Jeddah and Madinah airports. The results show variation of the number of initial infectious depending on the starting date. When initiating the infections within the peak weeks of arrivals, the initial size increased due to the increase in the number of arriving flights.

5.2 Epidemic Spread at Airport

The arriving pilgrims at Jeddah and Madinah airports were seeded with one infected individual per flight starting from the first day of arrival. Since the

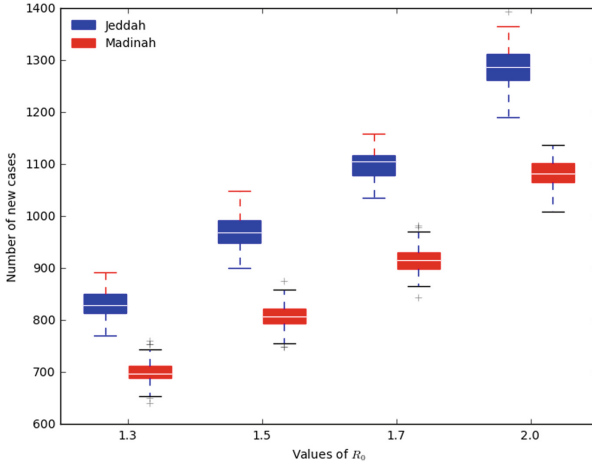


Fig. 3. Boxplots showing the distributions of the number of new cases (exposed individuals) at the end of arrival phase at Jeddah and Madinah airports, according to different values of R_0 (1.3, 1.5, 1.7 and 2.0).

total duration of mixing at the airport lasts for a few hours (2–3 h on average), which is shorter than the latent and infectious periods of the epidemic, there will be no changes in the disease status of the initial infected pilgrims. The only transition that can be seen after running the epidemic model is in the number of exposed individuals as susceptible become exposed when being infected by an infectious individual. As shown in Fig. 3, with the value of $R_0 = 1.3$, the maximum number of new cases reaches 0.2% of the total susceptibles. For $R_0 = 1.5$ and $R_0 = 1.7$, the average number of new cases ranges from 0.22% to 0.25%. For $R_0 = 2.0$, the total number of newly infected pilgrims is 0.33%. These results are obtained in the absence of any healthy behavior (e.g., wearing a mask) or control measures (e.g., vaccination). Random assignment of healthy behaviors to 50% of the arriving infectious individuals reduced the number of cases approximately by half.

5.3 Epidemic Spread at Makkah and Madinah

After executing the epidemic model for the airport arrival phase, the mixing patterns and the disease transmission among the pilgrims are simulated at the holy cities of Makkah and Madinah. We run the epidemic model at five time slots per day to reflect the five prayers performed in the grand mosque in each city. For each infectious individual, contacts from the current susceptible pilgrims are generated within each time slot based on the demographic of that individual. At the end of each day, the health timer for each pilgrim in the current population is updated and the health status changed accordingly. As shown in Fig. 4(a), the number of pilgrims in Makkah increased on a daily basis by the international pilgrims arriving from Jeddah airport. As described in Sect. 2, pilgrims arriving

via Madinah airport spend few days in the city then move to Makkah. Therefore, as shown in Fig. 4(c), the number of pilgrims in Madinah increased by new arrivals and decreased by how many pilgrims departed to Makkah. Changes in the number of susceptibles in each city as illustrated in Fig. 4(a) and (c) followed the same pattern of pilgrims arrival. However, the decline observed in the number of susceptibles on the last two days in Makkah resulted from successful disease transmission to a large number of individuals being in contact with all the accumulated infectious pilgrims in the final days of arrival.

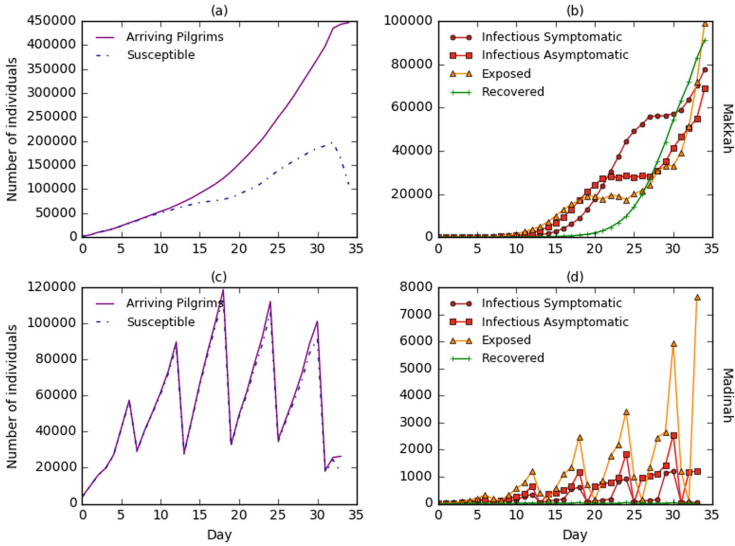


Fig. 4. The daily arrival patterns and results of the epidemic model within the 35 days of arrival in Makkah and Madinah. Figure (a) and (c) show the size of the arriving population along with changes in the susceptible compartment in Makkah and Madinah respectively and the results of the epidemic model are shown figure (b) and (d) with value of $R_0 = 1.3$.

The epidemic model is simulated in both cities with the infectious timer for the arriving infected individuals set to the first day of the infectious period. This scenario provides infectious individuals with a window of 7 to 10 days to mix and infect others. Following this extreme scenario and without applying any healthy behavior or implementing intervention measures, we expected to see a high number of new cases. However, the overall number of exposed individuals is less than the number of infectious pilgrims as the latent period lasts only for four days, at that time, individuals in the exposed compartment transfer to being infectious asymptomatic. In Fig. 4(b) and (d) show the daily incidence of exposed, infected and recovered pilgrims for the entire arrival period in Makkah and Madinah respectively. At the end of the arrival phase in Makkah, the epidemic model predict an average of 20% exposed and 30% infectious individuals

of the total pilgrims with 22% recovered individuals who were removed from the infectious compartment. While these numbers increase in Makkah due to the accumulation of pilgrims on a daily basis, the number of exposed individuals in Madinah fluctuates with several peaks during the days with higher numbers of arriving pilgrims. As illustrated in Fig. 4(d), the highest number of infectious individuals reaches 4% of the current pilgrims in Madinah on the last week of arrival, and the number of exposed pilgrims at the end of the arrival period is 29%.

The results obtained from our simulation experiments identified the impact of several factors on the progression of an epidemic at the arrival stage of a global event. These factors include the initial size of the epidemic, population density, and duration of the arrival phase compared to the length of the infectious and latent periods of the disease to be modeled. Diseases with infectious period less than the arrival stage duration may lead to increases in new cases who then have a potential to interact and infect other participants throughout the different stages of the mass gathering.

6 Conclusion

Major global mass gatherings require advanced planning and preparedness for public health hazards. The risk of emerging pandemic of infectious diseases within the timing of these events might hinder participants from attending or prevent many countries from sending their teams or missions. Epidemic modeling at the arrival stage of global mass gatherings can be applied proactively to provide initial indicators of the potential epidemics at these events. The results of this study highlight the impact of the initial number of infected participants, population density, and heterogeneous contact patterns on disease transmission at the arrival stage of Hajj. The model discussed in this study can be used to simulate the epidemic of influenza H1N1 at the different rituals of Hajj up to the departure of international participants, and the final estimates of infected individuals can be validated against results obtained from several published studies tracking H1N1 pandemics among returning international participants and their contacts after Hajj season. The proposed model can be used to study spread of infectious diseases in different mass gatherings by changing the duration of the arrival period, the travel patterns, and demographics of the expected arriving population. The epidemic model can help to explore different measures to prevent and control potential epidemics before the starting of a mass gathering.

References

1. Abubakar, I., Gautret, P., Brunette, G.W., Blumberg, L., Johnson, D., Pomeroy, G., Memish, Z.A., Barbeschi, M., Khan, A.S.: Global perspectives for prevention of infectious diseases associated with mass gatherings. *Lancet. Infect. Dis.* **12**(1), 66–74 (2012)

2. Ajelli, M., Gonçalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J.J., Merler, S., Vespignani, A.: Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC Infect. Dis.* **10**(1), 190 (2010)
3. Al-Tawfiq, J.A., Gautret, P., Benkouiten, S., Memish, Z.A.: Mass gatherings and the spread of respiratory infections. Lessons from the Hajj. *Ann. Am. Thorac. Soc.* **13**(6), 759–765 (2016)
4. Balcan, D., Gonçalves, B., Hu, H., Ramasco, J.J., Colizza, V., Vespignani, A.: Modeling the spatial spread of infectious diseases: the global epidemic and mobility computational model. *J. Comput. Sci.* **1**(3), 132–145 (2010)
5. Boëlle, P.Y., Ansart, S., Cori, A., Valleron, A.J.: Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: a review. *Influenza Other Respir. Viruses* **5**(5), 306–316 (2011)
6. Carley, K.M., Fridsma, D.B., Casman, E., Yahja, A., Altman, N., Chen, L.C., Kaminsky, B., Nave, D.: BioWar: scalable agent-based model of bioattacks. *IEEE Trans. Syst. Man Cybern.-Part A: Syst. Hum.* **36**(2), 252–265 (2006)
7. Chao, D.L., Halloran, M.E., Obenchain, V.J., Longini Jr., I.M.: FluTE, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput. Biol.* **6**(1), e1000656 (2010)
8. Chowell, G., Nishiura, H., Viboud, C.: Modeling rapidly disseminating infectious disease during mass gatherings. *BMC Med.* **10**(1), 159 (2012)
9. Coburn, B.J., Wagner, B.G., Blower, S.: Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC Med.* **7**(1), 30 (2009)
10. Dridi, M.H.: Tracking individual targets in high density crowd scenes analysis of a video recording in Hajj 2009. *Curr. Urban Stud.* **3**, 35–53 (2015)
11. Fiore, A.E., Fry, A., Shay, D., Gubareva, L., Bresee, J.S., Uyeki, T.M., Centers for Disease Control and Prevention (CDC), et al.: Antiviral agents for the treatment and chemoprophylaxis of influenza: recommendations of the advisory committee on immunization practices (ACIP). *MMWR Recomm. Rep.* **60**(1), 1–24 (2011)
12. Fraser, C., Donnelly, C.A., Cauchemez, S., Hanage, W.P., Van Kerkhove, M.D., Hollingsworth, T.D., Griffin, J., Baggaley, R.F., Jenkins, H.E., Lyons, E.J., et al.: Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* **324**(5934), 1557–1561 (2009)
13. Hu, H., Nigmatulina, K., Eckhoff, P.: The scaling of contact rates with population density for the infectious disease models. *Math. Biosci.* **244**(2), 125–134 (2013)
14. Khan, K., Arino, J., Hu, W., Raposo, P., Sears, J., Calderon, F., Heidebrecht, C., Macdonald, M., Liauw, J., Chan, A., et al.: Spread of a novel influenza A (H1N1) virus via global airline transportation. *N. Engl. J. Med.* **361**(2), 212–214 (2009)
15. Memish, Z.A., Zumla, A., Alhakeem, R.F., Assiri, A., Turkestani, A., Al Harby, K.D., Alyemni, M., Dhafar, K., Gautret, P., Barbeschi, M., et al.: Hajj: infectious disease surveillance and control. *Lancet* **383**(9934), 2073–2082 (2014)
16. Morse, S.S.: Factors in the emergence of infectious diseases. In: Price-Smith, A.T. (ed.) *Plagues and Politics*. GIS, pp. 8–26. Palgrave Macmillan UK, London (2001). <https://doi.org/10.1057/9780230524248.2>
17. Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G.S., Wallinga, J., et al.: Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**(3), e74 (2008)
18. Patlolla, P., Gunupudi, V., Mikler, A.R., Jacob, R.T.: Agent-based simulation tools in computational epidemiology. In: Böhme, T., Larios Rosillo, V.M., Unger, H., Unger, H. (eds.) *IICS 2004*. LNCS, vol. 3473, pp. 212–223. Springer, Heidelberg (2006). https://doi.org/10.1007/11553762_21

19. Perez, L., Dragicevic, S.: An agent-based approach for modeling dynamics of contagious disease spread. *Int. J. Health Geogr.* **8**(1), 50 (2009)
20. Shi, P., Keskinocak, P., Swann, J.L., Lee, B.Y.: The impact of mass gatherings and holiday traveling on the course of an influenza pandemic: a computational model. *BMC Public Health* **10**(1), 778 (2010)
21. Steffen, R., Bouchama, A., Johansson, A., Dvorak, J., Isla, N., Smallwood, C., Memish, Z.A.: Non-communicable health risks during mass gatherings. *Lancet. Infect. Dis.* **12**(2), 142–149 (2012)
22. Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Colizza, V., Isella, L., Régis, C., Pinton, J.F., Khanafer, N., Van den Broeck, W., et al.: Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Med.* **9**(1), 87 (2011)
23. Tabatabaei, S.M., Metanat, M., et al.: Mass gatherings and infectious diseases epidemiology and surveillance. *Int. J. Infect.* **2**(2) (2015)
24. Tanaka, G., Urabe, C., Aihara, K.: Random and targeted interventions for epidemic control in metapopulation models. *Sci. Rep.* **4**, 5522 (2014)
25. Tian, D., Liu, C., Sheng, Z., Chen, M., Wang, Y.: Analytical model of spread of epidemics in open finite regions. *IEEE Access* **5**, 9673–9681 (2017)
26. Tuite, A.R., Greer, A.L., Whelan, M., Winter, A.L., Lee, B., Yan, P., Wu, J., Moghadas, S., Buckeridge, D., Pourbohloul, B., et al.: Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *Can. Med. Assoc. J.* **182**(2), 131–136 (2010)
27. Yang, Y., Sugimoto, J.D., Halloran, M.E., Basta, N.E., Chao, D.L., Matrajt, L., Potter, G., Kenah, E., Longini, I.M.: The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* **326**(5953), 729–733 (2009)



Exploring In-Game Reward Mechanisms in Diaquarium – A Serious Game for Children with Type 1 Diabetes

Ida Charlotte Rønningen¹, Eirik Årsand², and Gunnar Hartvigsen¹(✉)

¹ Department of Computer Science,
University of Tromsø - The Arctic University of Norway, Tromsø, Norway
ida.charlotte.ronningen@gmail.com,
gunnar.hartvigsen@uit.no

² Norwegian Centre for E-health Research,
University Hospital of North Norway, Tromsø, Norway
eirik.arsand@ehealthresearch.no

Abstract. When developing serious games for health, the main goal is to use game mechanisms in a way that the users decide to extend their playing time, complete all levels within the game, and thereby gain progression and intended learning with regard to disease management. One major concern when developing games for health is, therefore, the possibility of users who withdraw from the game before completed. A game, with a rapid descending popularity and users quitting gameplay early, fails to provide medical education to patients and is thus useless. For that reason, motivational game elements, such as in-game rewards, have been heavily used when designing serious games. This paper identifies and suggests several reinforcement mechanisms within serious games and explores how they can be applied in diabetes. The game called Diaquarium, a serious game for children with Type 1 diabetes, provides knowledge regarding how nutrition, blood glucose levels, and insulin interplay for this patient group. A prototype has been developed to demonstrate its concept and some game mechanisms with help of Unity 3D game engine and the C# programming language. Game design, requirements and suggestions for the project, were gathered through literature review, attending workshops, meetings and discussions with experts, as well as feedback from a related user group through a questionnaire. The questionnaire was distributed to an elementary school class, involving nine 9-year-old children. The questionnaire examined and collected feedback regarding the game outline, usability, and preferred reward mechanisms in the Diaquarium game. Despite a short period of testing and a limited test group with non-diabetic children, the game was recognized as attractive and moderately difficult within the potential user group. The analysis suggests that rewards are highly a matter of preference. Simultaneously, there were indications that some of the rewards were more favorable than others. It appears that rewards serving a purpose within the game, e.g., potentially effect progression in the gameplay, are more favorable than the opposite rewards serving no purpose. The findings were highly valued and taken into consideration during the design process of exploring the in-game rewards of the Diaquarium.

Keywords: Serious games · Diabetes · Learning · Motivation
Reward mechanisms

1 Introduction

Learning to manage Type 1 diabetes (T1D) as a child can be a complicated adjustment, handling both medical and physiological changes in an already challenging stage of life caused by rapidly changing body functions. T1D is hard to regulate, and with the stressful situations all children experiences, T1D can create a whole new level of confusion and worry for the child when it comes to disease management [1]. Examples include school and social pressures (especially from trying to fit in), leisure activities, illness, and travels. Children also have difficulties understanding the severity of health complications that can occur if the disease is not correctly supervised.

Serious games often require that the users complete all levels to receive a full education. Thus, developers need to address the possibility of users quitting the game before finished off. Motivational game techniques and mechanisms are well known in the game industry, yet some serious games are useless if dismissed too early and thereby failing to distribute knowledge because they were perceived as boring [2]. Learning how to master motivation and engagement in serious games is, therefore, essential and valuable if one is aiming for increased knowledge in disease management through gameplay [3]. Reward mechanisms in games are known to have the potential to increase motivation and achieve learning, but research on differences in reward techniques and how their various characteristics can be applied in serious games and thereby impact users in regard to their likelihood to play again seem rather limited.

In light of the diabetes self-management difficulties for children addressed above, serious games have the potential to boost patients' medical treatment program in a positive manner. However, some of these games do not consider motivational in-game mechanism [2], which opens opportunities for consumers to quit the game earlier than intended and miss important knowledge. Göbel et al. [4] state that the reason why serious games are still missing market breakthrough is, among others, caused by poor quality of existing serious games. Children who are recently diagnosed with T1D, along with their families and friends, most often start at a zero-knowledge base. Therefore, it is important that we motivate them to not withdraw from playing a game that is designed to increase their self-management skills, skills that may be essential for their future health and to prevent possible consequences of incorrect care of diabetes.

The goal of this MSc project was to develop an educational game for children diagnosed with T1D, concerning suitable in-game reward techniques [5]. (The project is part of a larger serious game project for children with T1D.) The primary goal has been to investigate and explore various reinforcement techniques that exists in games and experiment how they can be applied in the best manner to increase patients' motivation to play. The main research problem was: *(Q) How can various reinforcement techniques be applied in serious games to increase play-act motivation for children with T1D?* This was divided into the following sub-problems related to reward and learning:

Q1: What types of rewards and reward techniques exist in serious games? Q2: What characterizes a good reward technique in serious games? Q3: How can rewards in serious games be applied in the best possible way? Q4: What are the most important tasks in self-management of T1D? Q5: How can T1D self-management tasks be presented in a serious game?

The paper discusses relevant theory related to reward a learning, presents the Diaquarium serious game for children with T1D and discusses the test results.

2 Methods

This project applied an engineering approach to problem solving that included the following steps: (1) User requirements and systems design: (1a) Early prototype (only addressing parts of the main game elements); (2) Data collection: (2a) Literature review, (2b) Discussions with experts and colleagues; (3) Experimentation: (3a) Online questionnaire for children, (3b) Short movie to illustrate the early prototype; and, (4) Evaluation.

Due to limited functionality of the prototype, a movie of the current prototype was made to better illustrate its functionality. Also, an anonymized questionnaire was set up for the test-group including the most important illustrations of the game concept, incorporating different game scenarios with various rewards. The feedback retrieved from the questionnaire provided valuable information concerning usability and preferred rewards the game, even though the game prototype was not complete. Finally, the results from the questionnaire were analyzed.

The game engine Unity 3D (www.unity3d.com) was used as the development platform for this project, enabling the developer to focus solely on the game logic and experimentation. In addition, Unity technologies provides their own full-features built-in editor, called MonoDevelop, that was used developing the project throughout. The C# programming language was used for scripting.

3 Rewards in Games

According to McKernan et al. [6], game scholars generally define a reward as “any game item or feature that reinforces particular in-game behaviors”, and that this form of behavior reinforcement may motivate players to continue game-play. Similarly, Phillips et al. [7] define rewards in games as “a positive return that serves to reinforce player behavior within a video game”. Schell [8] described rewards as a type of feedback, or “the way the game tells the player ‘you have done well’”. Rewards have three fundamental functions on individuals; (i) they evoke learning as they make individuals come back for more, (ii) they induce approach and consummator behavior for acquire the reward object, and, (iii) they stimulate subjective feelings of pleasure and induce positive emotions [9, 10]. Howard-Jones and Jay [11] studied the link between reward and learning in the context of reinforcement learning. They claim that the term “reward” differs in meaning within circumstances of education and cognitive neuroscience.

There are many different types of rewards in games, but there has been limited empirical research with regard to the classification and types of video game rewards [7]. We have identified six papers as the most influential in this respect. These are summarized below:

- Hallford and Hallford [12]: Rewards of glory; Rewards of sustenance; Rewards of access; and, Rewards of facility.
- Oxland [13]: Game-play rewards; Hidden rewards and secrets; Impetus rewards; and, Visual rewards.
- King et al. [14]: Praise; Points; Prolonged play; A gateway; Spectacle; Expression; Powers; Resources; and, Completion.
- Wang and Sun [15]: General reward types; Meta-game reward features; Intermittent reward features; and, Negative Reward features.
- Chou [16]: Score reward system; Experience point reward system; Item granting system rewards; Resources; Achievement systems; Feedback messages; Plot animation and pictures; and, Unlocking mechanisms.
- Johnson [17]: Fixed Action Rewards (earned lunch); Random Rewards (mystery box); Sudden Rewards (easter eggs); Rolling Rewards (lottery); Social Treasure (gift from friends); and, Pricing Pacing (small pieces).

4 Game Scenario

The game starts with a goldfish named Otis, who presents the story behind Diaquarium. Otis reports that he is diagnosed with T1D, and that he used to live in a goldfish shop. Unfortunately, the goldfish shop did not provide him with the correct food or medicine in order for him to stay healthy. One day, someone from Diaquarium adopted him, and declared that Otis is going to live temporarily in their aquarium until he finds his final home. In Diaquarium, the goldfishes living there have all T1D, and those who work in this profession are providing correct health care to prevent later health consequences. Now, Diaquarium has a lot to do, and they are looking for an assistant that can help them out with the meal and medicine duties. When the player agrees to help, the game begin.

There are no levels in Diaquarium, a design choice made early in the process. The intention was to create a never-ending game (levels would eventually have created an end), so that player input (how active they are (clicks) as well as length in gameplay) could be correctly measured. The idea to measure playtime length was discarded, but the concept with no levels remained. Even though there are no levels in the Diaquarium, players have the opportunity to increase the number of goldfishes as they like if they have enough resources to do so, which will increase the difficulty of the game.

When a player clicks on a goldfish in Diaquarium, some main information reveals itself in respect to that specific goldfish, such as name, health status according to a glucose barometer, food-button, and potentially a new-home-button if the goldfish is ready to move to a final home. The player must choose combinations of food and correct amount of insulin according to the glucose status to keep the goldfish healthy

and satisfied. To continue playing, players must choose to provide new homes to the goldfishes, as well as buying new goldfishes from the goldfish shop.

5 Game Outline

When the game starts, the main character Otis appears on the screen and reveals the background story too the player through speech bobbles (Fig. 1). After making the player familiar with the background story, the aquarium covers the screen and the user is ready to play the assistant role in Diaquarium. First, Otis demonstrates the main features of how the gameplay is carried out by some small animations and speech bobbles. These features are known as the core game mechanisms, also identified as the game outline.

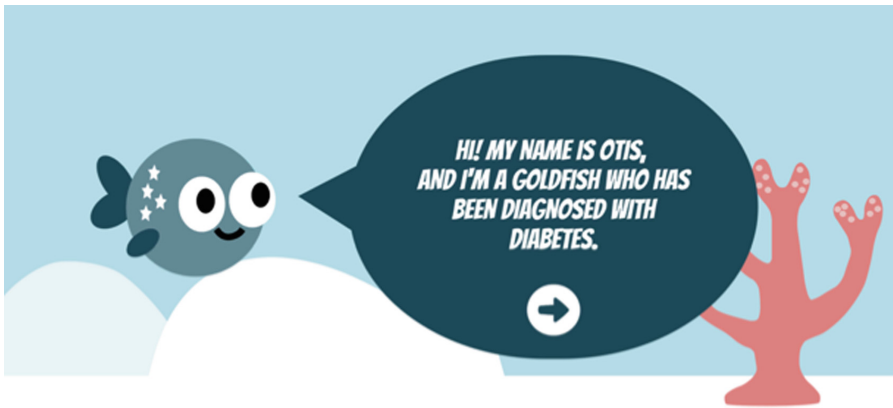


Fig. 1. Game start in the Diaquarium.

Otis shows the glucose barometer as an imprint of his current glucose level (health) at any time. If the blood glucose level drops low (yellow zone in the glucose barometer), or too high (red zone in the glucose barometer), the goldfish facial expression changes (Fig. 2). In order to keep the blood glucose level normal (green zone), the player must provide correct food and insulin dosage. When the player interacts with the food-button, a new window appears sliding in from the left side of the screen (Fig. 3).

Next, Otis points out how the assistant must drag-and-drop different food-items on the plate and choose insulin dosage in order to complete the task. The amount of insulin affects the blood glucose levels according to the food items chosen. When goldfishes move to their new homes, they will naturally be removed from the Diaquarium and do not require any more assistance. For the sake of continuous play and increasing difficulty of the game, the goldfish shows the player the possibility of visiting the local goldfish and save some more goldfishes.

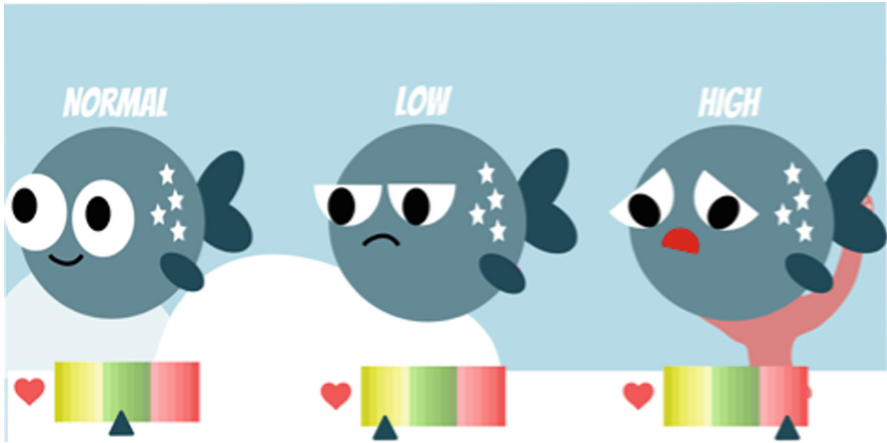


Fig. 2. Emotions in relation to glucose level in bloodstream. (Color figure online)

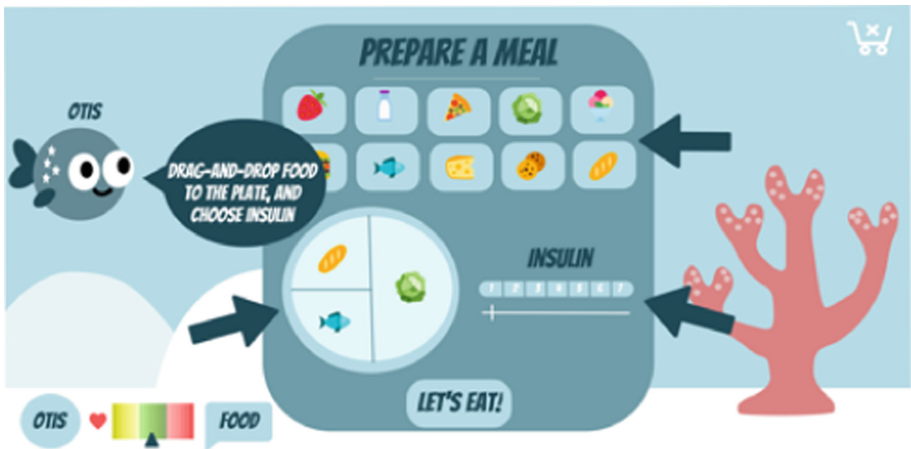


Fig. 3. Prepare a meal. (Color figure online)

6 Rewards Applied in the Diaquarium

6.1 Positive Rewards

The player receives his/her first reward when he/she starts to play. The reward is a surprise, and the player will never know what he/she gets until he/she has launched the game. The reward can be any of the following: Gold coins; A free goldfish; or, A medicine packet that keeps the goldfish within the green zone on the health bar for approximately 1 min (e.g., the player does not need to do anything for that goldfish within this amount of time).

Another reward is praise for checking on a particular goldfish. The player must click on the goldfish to check its health status, and the game randomly makes that a particular goldfish talks to the player. The goldfish tells the player in different ways how good he/she is doing with the assistant job. Other praises could possibly be; “Amazing! You are doing a really good job!”, “Wonderful! You are so thoughtful”, “You take good care of me, thank you!”.

The player (assistant) will be rewarded if he/she manages to stay within the green zone of the health line for 1 min at a time (Fig. 4). Whenever a goldfish enters the green zone, a clock starts counting down from 1 min. Note that each goldfish has its own exclusive clock for each respective health line. If the goldfish moves over to the yellow zone or red zone of their health line, the clock will disappear for that particular goldfish. The clock does not cache timestamps when a goldfish leaves the green zone, and will therefore always start over again when the green zone is entered. Whenever the count-down is completed, the player receives 100 gold coins as well as praise for doing a good job. The game is fair, and the clock starts immediately a new count-down when 1 min has been completed and the goldfish is still staying inside the green zone.

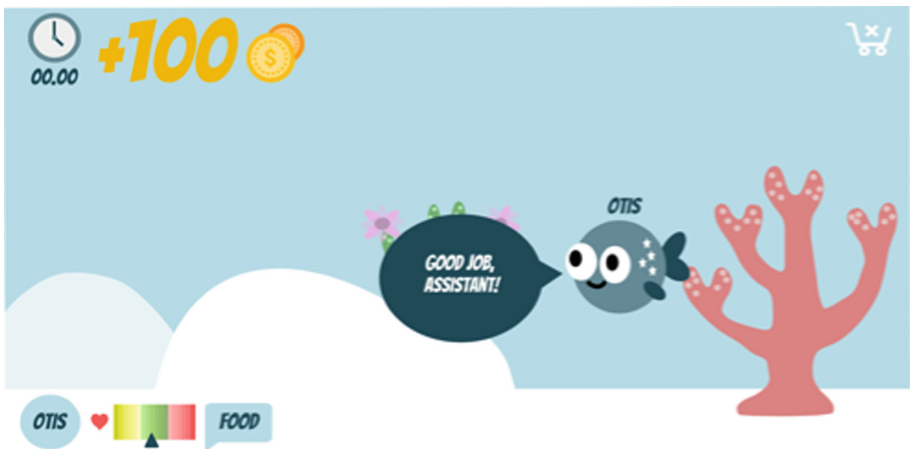


Fig. 4. Illustration of reward that is received after 1 min of staying within the green zone of the health line. (Color figure online)

If the player manages to stay within the green-zone for 3 count-downs in a row, the player is being rewarded with a medicine packet. The medicine packet pauses the goldfish within the green-zone for 1 min, making the player earn free gold coins as the clock continues to tick down.

If the goldfish receives correct food and insulin dosages, it will express its gratitude with words such as “Yum!”, “I like this!”, or “You’re a good chef!”. If the player carries out three correct meals in a row with appropriate insulin dosages, a reward will be earned through a price in the lucky wheel (Fig. 5). The lucky wheel contains four types of rewards: (A) Slow increase or decrease in blood glucose for 1 min; (B) A medicine packet that keeps the goldfish within the green zone on the health bar for



Fig. 5. Lucky wheel with rewards. (Color figure online)

approximately 1 min (e.g., the player does not need to do anything for that goldfish within this amount of time); (C) Praise; and, (D) Gold coins.

When the goldfish is ready to move to its final home, the player gets rewarded with gold coins. However, there is a little twist with this particular reward – the player never actually knows how much he/she will get paid until he/she gets it. The player can, e.g., earn 100 gold coins one time, and 50 gold coins the next time. In addition, the player is rewarded with a new type of goldfish in the goldfish shop for every random number in range of 1 to 5 goldfishes who found their final home. To exemplify, the player will be rewarded with a new type of goldfish in the goldfish shop after deployment of 3 goldfishes (not necessarily in a row), and next time after the deployment of 1 goldfish.

6.2 Negative Reinforcement

The game scenario of the Diaquarium includes some negative reinforcement as well. If the goldfish experiences long time-periods of too low or too high blood glucose levels, extra seconds will be added to the goldfish's time in Diaquarium before it finds its final home. The player must therefore do some kind of planning along with actions to keep the goldfish within normal blood glucose levels so it can move to its new home and the player can rescue new goldfishes, and thereby continue to play.

Also, too long time-periods of too low or too high blood glucose levels will result in having the goldfish transferred to the hospital, e.g., removed from Diaquarium without the player's consent. Besides, the player must pay for the hospital stay, losing some of the already earned rewards (Fig. 6). Additionally, if players deliver wrong food regarding the situation and/or inappropriate amount of insulin, the goldfish will

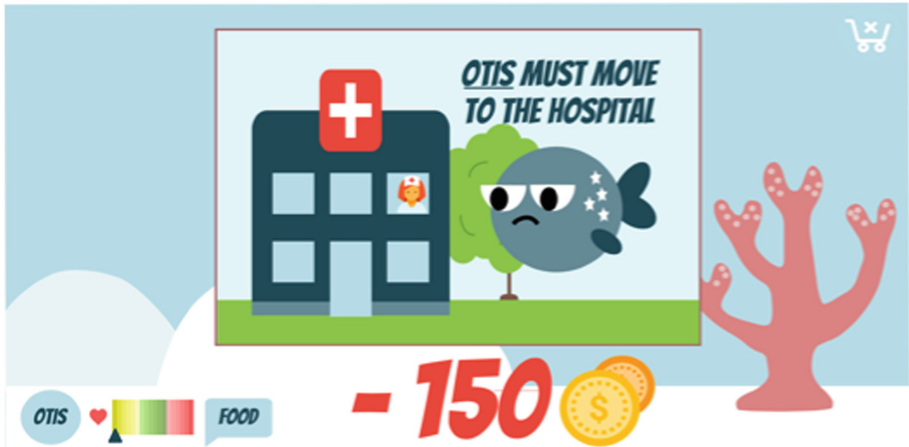


Fig. 6. Pay for hospital stay.

respond with quotations like “I don’t like this!” or “This was not good for me....” along with adjustments in blood glucose levels.

7 Discussion

The online questionnaire was distributed to a class of nine 9-year-old pupils. Most pupils answered that they liked goldfishes, and there was $n = 0$ participants who didn’t like goldfishes at all. When asked what type of rewards they would like to receive when they managed to provide correct food and medicine to the goldfish, half of all the pupils wanted to receive gold coins that they could use to buy new goldfishes. This was as expected, as gold coins were represented as the only resource for progression in the gameplay (e.g., the only reward that could buy goldfishes). The rest of the pupils wanted either praise or a surprise to honor the correct action accomplished. Praise is a well-known attribute in in-game rewards, and this answer correlate with children’s need for recognition and explicit assurance that they managed to do the correct action. It is important to clarify that the concept of praise in this project is addressed as a type of feedback and acknowledgement. For those pupils who answered surprise as the desired reward, it suggests that the excitement when considering an unknown reward is so strong that they would potentially risk getting (what they would consider) a bad reward for the thrill of it. According to Johnson [17], elements of surprise and unexpected rewards keep players engaged, pushing them through a tough level and lead them to mastery. This is confirmed in a later question, where $n = 9$ would rather be rewarded spinning a lucky wheel to see what reward they will get, instead of a static reward of gold coins, which would have been a safer choice regarding progression in gameplay.

It is interesting to see that for the next question, regarding what reward they would like the most, $n = 2$ of the participants decided to answer a different type of reward.

They had both answered gold coins as the reward they liked the most in the previous question, but now, $n = 1$ of participant went for the surprise reward, and $n = 1$ wanted a medicine packet to keep the goldfish healthy for 3 min. These answers exemplify one of the biggest difficulties game designers and developers faces: users' decisions and opinions regarding what they like can immediately change due to their personal preferences. This specific challenge for game designers and developers is also evident when comparing the least wanted rewards. Most of the participants answered music as the most minimal reward of all rewards presented, but some of them also answered points used in a high score list and praise. Recap to the previous answers, praise was one of the most wanted reward for some of the participants.

There is only one question in the questionnaire regarding reward schedules. It asks how they want to receive rewards when they have managed to stay within the green zone of the glucose barometer. The plurality of the participants answered that they preferred static reward (e.g., reward after 5 min within green zone). This answer contradicts with the reward schedule literature that claims that the variable interval reward schedule is more attractive than the fixed interval reward schedule, because the user never knows when the next reward arrives. According to Sylvester [18], fixed reward ratio schedules are poor motivators, because they encourage long periods of inactivity, waiting for the next reward. It is in these long periods of inactivity that players easily get up and walk away from the game [18]. Variable reward ratio schedules, on the other hand, is the most powerful simple motivator of all schedules [18], because it will keep players playing. Since there always will be a chance to get a reward the next time, players will continue gameplay, hoping the next enemy will drop the big payoff [18].

We then asked what they thought about negative reinforcement in light of having to pay for the goldfish's hospital stay. To our surprise, $n = 8$ of participants chose answers that articulated OK. This result complies with the literature that suggests that negative reinforcement does not lead to negative emotions intermittently. Instead, the players respond to failures with excitement, interest, and joy, getting highly motivated to return to the task being optimistic about reaching their goals [19].

One question investigated the social aspect of Diaquarium that was considered to be integrated as a component within the game for future work. Even though it is not part of the current design, we wanted to see if the children liked the concept where they could be able to move to other aquariums and chat with other goldfishes. Of all participants answering the question, $n = 7$ were positive to the idea, which suggests that the Diaquarium surely has the potential to have some sort of social components included in the game outline. Social aspects are considered to exist as a functionality of serious games that can enhance play personalization and interactivity [20].

The answers regarding usability indicated that the majority of participants think that the game look not too difficult, and that they mostly understood how they could achieve rewards. However, $n = 1$ answered that he thought the game look difficult and that he did not understood how he could receive rewards.

8 Concluding Remarks

The main research problem in the overall project, which this project is part of, was to explore in-game rewards in serious games, and apply these techniques in an educational game that was designed and implemented to help children with T1D to improve their self-management skills. The main goal of the project reported in this paper was to identify various reward techniques used in games and address how they could be applied to prevent children from quitting gameplay, and consequently lose significant information and knowledge related to diabetes. Serious games with characteristics to be fully engaging and motivational for children can likely help them to avoid disease complications in the future. As a result, the Diaquarium was designed to explore these kinds of problems. Also, an early prototype was implemented to illustrate the main characteristics of the game itself.

The design of the Diaquarium is achieved by identifications of significance in-game rewards, game features, and psychological behavior regarding learning and motivation. Also, feedback and suggestions from a potential user group, professors in field of game development, professors in field of psychology, experts in Diabetes Mellitus, as well as knowledge from workshops, were considered during the process of designing the diabetes.

Diabetes-related knowledge was integrated into the Diaquarium gameplay, where users of the application enhance their knowledge in respect to interaction regarding food, insulin, emotions, and blood glucose levels for individuals with the disease through the main story of the game. In the final stage of the project, a questionnaire was distributed to a class of nine ($n = 9$) 9-year-old children. The outcome from the questionnaire implies that least or most wanted rewards often are results of personal preference, but the result also suggests that some of the rewards presented was more/less wanted than others within the whole group of potential users. The most wanted rewards were considered a resource within the game as well as being surprising to the users. The least wanted reward was sound/music, where $n = 6$ agreed that this was not a reward they highly desired.

The findings regarding the sub-questions are [5]:

Q1: What types of rewards and reward techniques exist in serious games? In total, 36 types/categories of in-game rewards and 6 in-game reward schedules were identified.

Q2: What characterizes a good reward technique in serious games? A good reward technique acknowledges appropriate reward types/categories that can be applied in the game story. Most wanted and least wanted rewards differs for age groups and cultures, which are characteristics that shall be addressed to seamlessly fit the system. A good reward technique assures that there are no gaps in motivation that could possibly create windows where players decides to stop playing [18]. Thus, good reward techniques shall run several rewards and reward schedules simultaneously, so that whenever one schedule reach motivation bottom, the others are at their motivation peak.

Q3: How can rewards in serious games be applied in the best possible way? Rewards in games shall reinforce behavior, actions and decisions made throughout

to increase users learning and knowledge, as well as increase motivation. Rewards should be applied after gameplay action, so that the users know exactly why they are being rewarded and therefore are able to determinate outcome of decisions made. According to test results, rewards that surprises users seems to be highly engaging.

Q4: What are the most important tasks in self-management of T1D? The most important tasks required in self-management of T1D were in this project identified as: Healthy eating; Being active; Monitoring blood glucose, blood pressure, urine ketones, and weight; Problem solving; Reducing risk; and, Health coping.

Q5: How can T1D self-management tasks be presented in a serious game? The Diaquarium demonstrates how healthy eating behavior and blood glucose monitoring can be applied in a serious game, which are some of the acknowledged important behaviors required in self-management of T1D. The game solved this problem by applying T1D to the main characters in the game, and let these characters demonstrate diabetes related behaviors, where self-treatment behaviors had to be performed in order to have progression in gameplay.

References

1. JDRF: Teen Toolkit. JDRF, New York (2013). <http://www.jdrf.org/wp-content/uploads/2013/10/JDRFTEENTOOLKIT.pdf>. Accessed 12 Dec 2016
2. Mitgutsch, K., Alvarado, N.: Purposeful by design? A serious game design assessment framework. In: Proceedings of the International Conference on the Foundations of Digital Games (FDG 2012), pp. 121–128. ACM, New York (2012)
3. Lewis, M.W.: Analysis of the roles of “serious games” in helping teach health-related knowledge and skills and in changing behavior. *J. Diabetes Sci. Technol.* **1**(6), 918–920 (2007)
4. Göbel, S., Hugo, O., Kickmeier-Rust, M., Egenfeldt-Nielsen, S.: Serious games—economic and legal issues. In: Dörner, R., Göbel, S., Effelsberg, W., Wiemeyer, J. (eds.) *Serious Games*, pp. 303–318. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40612-1_11
5. Rønningen, I.C.: Exploring in-game rewards in the Diaquarium: a serious game for children with type 1 diabetes mellitus. Master’s thesis in Computer Science. University of Tromsø – The Arctic University of Norway, December 2016
6. McKernan, B., Martey, R.M., Stromer-Galley, J., et al.: We don’t need no stinkin’ badges: the impact of reward features and feeling rewarded in educational games. *Comput. Hum. Behav.* **45**, 299–306 (2015)
7. Phillips, C., Johnson, D., Wyeth, P.: Video game reward types. In: Proceedings of First International Conference on Gameful Design, Research, and Applications, pp. 103–106. ACM, New York (2013)
8. Schell, J.: *The Art of Game Design: A Book of Lenses*. Elsevier, Amsterdam (2008)
9. Schultz, W.: Multiple reward signals in the brain. *Nat. Rev. Neurosci.* **1**(3), 199–207 (2000)
10. Schultz, W.: Neural coding of basis reward terms of animal learning theory, game theory, microeconomics and behavioral ecology. *Curr. Opin. Neurobiol.* **14**(2), 139–147 (2004)
11. Howard-Jones, P., Jay, T.: Reward, learning and games. *Curr. Opin. Behav. Sci.* **10**, 65–72 (2016)

12. Hallford, N., Hallford, J.: *Swords and Circuitry: A Designer's Guide to Computer Role Playing Games*. Prime Publishing, Roseville (2001)
13. Oxland, K.: *Gameplay and Design*. Addison Wesley, Harlow (2004)
14. King, D.L., Delfabbro, P., Griffiths, M.D.: Video game structural characteristics: a new psychological taxonomy. *Int. J. Mental Health Addict.* **8**(1), 90–106 (2009)
15. Wang, H., Sun, C.-T.: Game reward systems: gaming experiences and social meanings. In: *Proceedings of the 2011 DiGRA International Conference: Think Design Play*. Utrecht School of the Arts (2011)
16. Chou, Y.: *The Six Contextual Types of Rewards in Gamification* (2013). <http://www.yukaichou.com/marketing-gamification/six-context-types-rewards-gamification>. Accessed 14 Dec 2016
17. Johnson, P.: *Surprise in Game Design* (2013). <https://www.sundoginteractive.com/blog/surprise-in-game-design>. Accessed 15 Dec 2016
18. Sylvester, T.: *Designing Games: A Guide to Engineering Experiences*. O'Reilly Media, Sebastopol (2013)
19. Granic, I., Lobel, A., Engels, R.C.M.E.: The benefits of playing video games. *Am. Psychol.* **69**(1), 66–78 (2014)
20. Konert, J., Göbel, S., Steinmetz, R.: Towards social serious games. In: Connolly, T., Felicia, P., Neville, G., Tabirca, S. (eds.) *Proceedings of the 6th European Conference in Game Based Learning (ECGBL)*, vol. 1, no. 1. Academic Bookshop, Cork, Ireland (2012)



An “Awareness” Environment for Clinical Decision Support in e-Health

Obinna Anya^{1(✉)} and Hissam Tawfik^{2(✉)}

¹ San Jose, CA, USA
o.anya@ieee.org

² School of Computing, Creative Technology and Engineering,
Leeds Beckett University, Caedmon 111, Headingley Campus,
Leeds LS6 3QS, UK
h.tawfik@leedsbeckett.ac.uk

Abstract. The notion of cross-boundary decision support has the potential to transform the design of future work environments for e-health through a connected system that allows for harnessing of healthcare information and expert knowledge across geographical boundaries for more effective decision-making. The trouble, however, is that the use of healthcare information in decision-making usually occurs within the context of a complex structure of clinical work practices that is often shaped by a wide range of factors, including organisational culture, local work contexts, socially constructed traditions of actions, experiences and patients’ circumstances. They vary across geographical and organisational boundaries, and have remained to date largely unaccounted for in the design of e-health systems. As a result, achieving the e-health vision of ‘open’ clinical decision support requires a rethinking of key clinical and organisational processes in a manner that accommodates clinical work practice as a fundamental part of how clinicians work and make decisions in real-world settings. Drawing on the theories of human activity system and situation awareness as well as the belief-desire-intention architecture in AI, this paper presents the design of an awareness environment for cross-boundary clinical decision support in e-health that takes account of the concept of work practice as a design requirement. The proposed system shows that incorporating practice information into the design of e-health systems enhances their usefulness for ‘open’ clinical decision support.

Keywords: Clinical work practice modelling
Cross-boundary knowledge sharing and decision support
Practice-centred awareness · e-Health · Activity system · Framework

1 Introduction

Awareness is conceptualised differently in relation to two different scenarios. The first is in a co-located setting, where participants work towards achieving a common goal, e.g. in a control room [1]. The other is in distributed settings, where participants are remotely attempting to collaborate via some kind of technological support, e.g. in cross-boundary clinical decision support in e-health. Vyas [2] notes that the two

scenarios require different treatments of the issue of awareness. Whereas the concept and process of constructing awareness appear trivial, and easily taken for granted, in co-located contexts [3–6] in distributed settings, the process has to be mediated, e.g. by technology or practice, or effectively signaled by users [3] in distributed settings. In order to design for meaningful cross-boundary decision support in the latter, it becomes imperative to conceptualise effectively what awareness means, to derive a mechanism for understanding the numerous aspects of what awareness construction entails, and to decide what aspects of the other’s work – both from theoretical and practical perspectives – that one *needs to be aware of*.

The idea of constructing and maintaining awareness of one another’s working context has been explored considerably in literature [2, 7, 8], and has been applied to a number of real-world settings in such areas as healthcare, driving and simulation, and control and operating rooms [1, 5]. Different forms of awareness have been proposed and explored, including collaboration awareness, peripheral awareness, background awareness, mutual awareness, passive awareness, workspace awareness, experience-focused awareness, activity-based awareness, situation awareness, and context-based awareness [2, 5, 6, 8–10]. While the proliferation of forms of awareness is a clear indication that the term “awareness” is found to be equivocal [11], the various guises under which the concept has been studied has, nonetheless, contributed to the design of computational models of awareness that are able to explicate the complex array of interdependent and highly contingent circumstances, conditions and actions that are required to understand things and events in the world. Several researchers, such as [1–3, 9, 11, 12], have reported a number of characteristics of awareness. An analysis of these works, particularly from a practice-centred perspective, suggests that awareness consists of the following basic conceptual features:

- Awareness is determined by work and workplace setting. It is knowledge of the state of the environment, environment being a temporally and spatially bound setting for people interacting within it. It is by the people’s experiences of this interaction.
- Awareness is dynamic, because the environment changes, and thus awareness needs to be kept up to date [12].
- Maintaining awareness is not the main target of tasks. It is necessary, but not enough. It enables smooth completion of tasks [12].
- Awareness has a socio-technical dimension, and is important to both social and technical research.
- Awareness is both implied in people’s work and an integrated aspect of their practice [2].

We adopt, in this work, a practice-centred approach to the design of awareness environments for cross-boundary decision support in e-health. Awareness is an attribute of practical action, and, as such, is constructed as a part of how clinicians work. As rightly noted by [1, p. 318] “the ways in which individuals accomplish awareness is inextricably embedded in the activities in which they are engaged, and the ways in which those activities necessarily entail particular practices and procedures”. Awareness is purposeful and implied in our activities and practices. It involves not only knowledge about the dynamic and task-oriented, easily observable interactions within a

spatially and temporally bounded space, but also knowledge about what influences those interactions and the end to which the interactions are directed. What others will become aware of about our activities is dependent upon the motive for seeking such awareness. Moreover, the awareness is equally dependent upon the activities in which we and the others are engaged in, our domain and organisational context, and how we conduct those activities (i.e. work practice).

Awareness captures the social and work contexts of use in CSCW, decision support and health information systems. It conveys a broad notion of “taking heed of [the] context” [11, p. 286] of something or event. Observational studies of complex and knowledge-based problem-solving domains [5, 13, 14], have suggested that users often employ practices and procedures that not only fall outside of formal work processes, but were also hardly predictable during system design [3]. It has been suggested that our system design approach need to take into account diverse coordinative practices through which work is routinely and seamlessly integrated [2, p. 16]. These coordinative practices differ from domain to domain, from locality to locality, and from context to context. And particularly for healthcare, a firm grounding of design approaches in the concept of practice is crucial for building effective work support systems [14, 15].

This paper presents the design of an awareness environment for cross-boundary clinical decision support in e-health that integrates the human activity system and situation awareness with the belief-desire-intention (BDI) architecture in AI [16]. This paper posits that a viable starting point for the construction of human-centred AI systems for clinical decision support, especially across boundaries of communities of practice in e-health is to explore the potential use of socio-technical theories that have quite effectively framed the development of human-centred systems in other fields of systems engineering, air traffic control, and HCI.

2 Modeling Practice-Centred Awareness

Our starting point in developing a practice-centred awareness (PCA) model for cross-boundary clinical decision support in distributed communities of practice is the integration of the human activity system and situation awareness as depicted in Fig. 1. We define PCA as *an understanding of other people’s local work contexts, problem-solving approaches, circumstances and task requirements, which include the ontological, spatio-temporal and situational factors that provide causal explanations for and influence how they utilise available resources, and contextualise plans and*

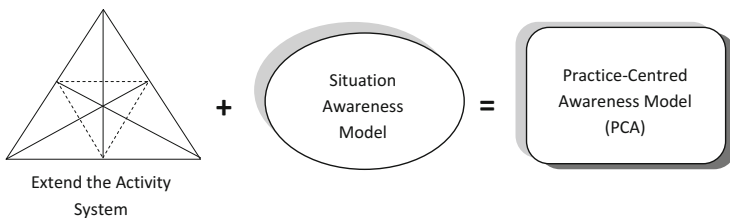


Fig. 1. Modeling practice-centred awareness

procedures to solve problems and achieve task goals in the real-world. Developing the PCA model involves three major steps. In the first step, we extend the basic model of Engeström’s activity system [17] in order to derive a *work practice model*. In the second step, we integrate the work practice model with Endsley’s SAW model [10]. Finally, we incorporate into that our proposed context model. One of the novelties of our modelling approach is its potential to provide a knowledge-oriented, ecology-level understanding of actions in a work system. By using the CHT approach, we emphasize the historicity and situatedness of actions as well as the interactivity required to perform work; by using the SAW, we are able to highlight the constant evolution of the knowledge and contextual variables required to work done.

We model work practice as context-driven interactions emerging out of a clinician’s engagement with the environment, drawing from their knowledge of the domain of medicine as well as every day socio-cultural understanding of clinical practice in their workplace. Our goal is to enable an enhanced understanding of clinical work contexts and problem requirements across distributed work settings for cross-boundary decision support. Hence, we define PCA as: *an understanding of other people’s local work contexts, problem-solving approaches, circumstances and task requirements, which include the ontological, spatio-temporal and situational factors that provide causal explanations for and influence how they utilise available resources, and contextualise plans and procedures to solve problems and achieve task goals in the real-world.*

3 Designing an Awareness Environment for Cross-Boundary Decision Support

Having introduced the general abstractions and mechanisms of the approach proposed in this work, we describe the design and development of the awareness environment proposed in this paper. We refer to the prototype as **Context-Aware cross-boundary clinical Decision support system in e-Health (CaDHealth)**. CaDHealth is designed as a *practice display system*. The content of the display is a visualisation of the PracticeFrame – a representation of located clinical problem-solving based on the ontological, stereotyped and situated work practice descriptions. The use of visualisation techniques or “displays” to represent knowledge [18], work contexts [5], expertise [19], and awareness information [7] at the interface level is not new. However, what is new, as far as we know, is the use of the approach to represent work processes at the work practice level. CaDHealth is developed as a frame-based representation of the situations and circumstances of a clinical work setting. To a large extent, the design was inspired by Gabbay and le May’s model of “clinical mindlines” [14] as a representation of the situated, internalised and practice-centred guidelines that serve as a clinician’s knowledge-in-practice-in-context (p. 101). As depicted in Fig. 2, CaDHealth supports users in two principal ways:

- Through the provision of Practice Display: This describes the *practice information*, which the systems provide to a collaborating expert in a remote work setting to enable them to gain knowledge (i.e. awareness) of the user’s work practice situation and patient’s needs in order to enable them to offer appropriate suggestions, and

- Through the provision of Enriched Decision Support Information Display: The suggestion provided by a collaborating expert is *enriched by the system* by infusing into it more information (e.g. from the system database) and *morphed for contextual adaptability* in order to provide the user with context-aware information (and enriched suggestion) to support their clinical decision.

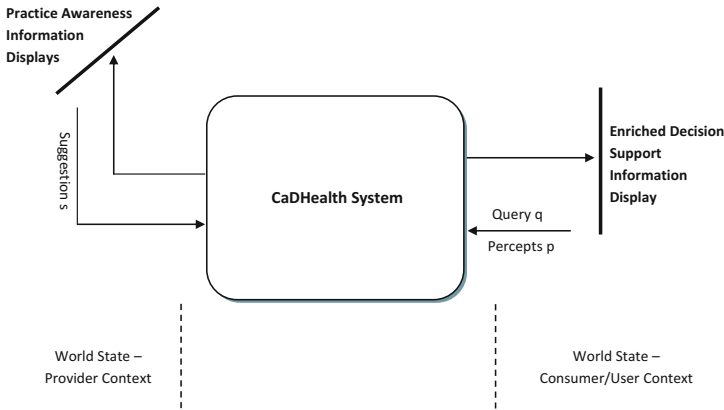


Fig. 2. CaDHealth showing two types of information provided for cross-boundary clinical decision support in e-health

3.1 CaDHealth Architecture

CaDHealth has been designed as an interactive practice information display system that allows clinicians to maintain awareness of work situations and problem requirements across work settings for cross-boundary decision support. From an HCI perspective, our task involves designing the system with the capability, (1) to enable awareness of work situation across regional and organisational boundaries, (2) to support clinical decision based on wide ranging factors that include cultural, circumstantial and interactional factors that influence how work actually gets done “in practice” as opposed to an abstract top-down functional model of an organisation’s work process, and (3) to avoid supplanting a clinician’s judgment. CaDHealth is designed as part of a hospital’s integrated e-health system¹ in order to portray the system’s ability to illuminate the role of stereotypes, situated and interactional factors in clinical decision making. Secondly, such approach is necessary for standardising practices across regions and organisations, and by incorporating guidelines and evidence-based information (as dampers), it ensures that local work practices are clinically compatible and that their independent application does not compromise patient safety [20].

¹ This was largely informed by the result of a user-centred study reported in [15, 23] where clinicians expressed preference for a cross-boundary decision support system that is integrated into their hospital’s health information system or broader e-health system.

The system architecture of CaDHealth is illustrated in Fig. 3. The system consists of three main components: the CaDHealth user interface (UI) layer, a cross-boundary collaboration layer and the PCA Manger. The UI and the cross-boundary collaboration layers are designed as client-side applications, whereas the practice-centred awareness (PCA) Manager, integrated into a HIS, constitutes the CaDHealth server. The backend includes a knowledge repository and the core system database. The knowledge repository stores domain models and practice models and percepts, i.e. the semantic, practice and perceptual memories respectively; the core system database is the working memory and stores clinical work processes and practice-aware decision models. In addition to these components, the infrastructure is potentially able to connect to external and cloud-based services, such as location-tracking services, sensors, actuators, RFID (Radio-Frequency Identification) readers, situation models as well as regional, organisational and domain-specific services. Because CaDHealth is integrated into a hospital’s larger HIS, the architecture includes a firewall, which ensures that (1) sensitive patient information, e.g. in patients’ health record or a hospital’s institution guideline, is *anonymised* before being used in cross-boundary decision support, and (2) only authorised and authenticated agents and services are granted access.

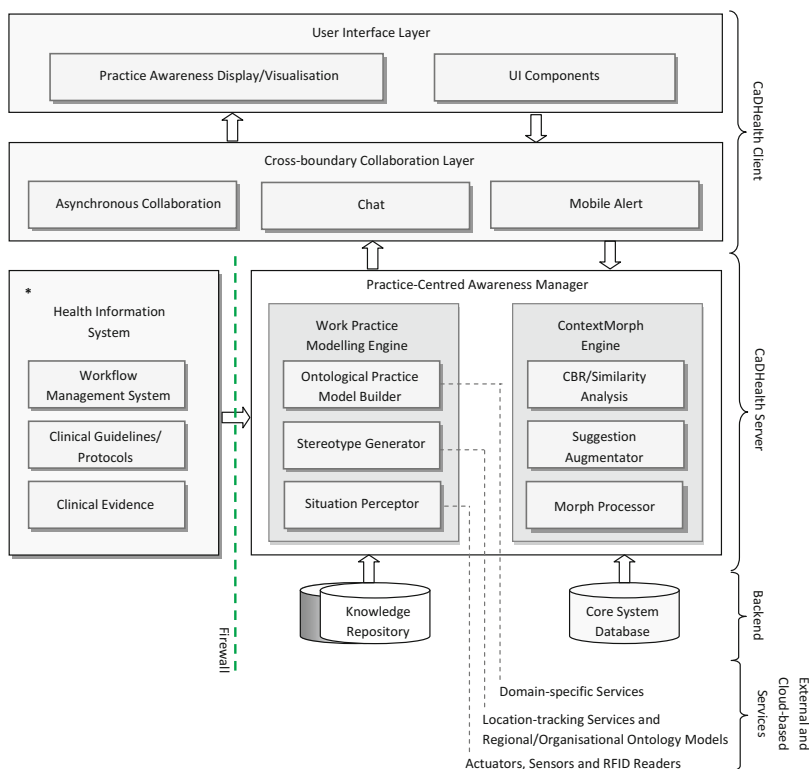


Fig. 3. The system architecture consisting of the user interface components, the cross-boundary collaboration layer, and the knowledge layer. The component labelled Health Information System* (HIS) denotes a hospital’s existing HIS and is not part of the CaDHealth system.

As shown in Fig. 3, CaDHealth is designed following a multi-tier system architecture in which the client, server and backend sides as well as required external services could reside physically on different nodes on a network. However, the CaDHealth infrastructure is logically a peer-to-peer distributed system. What this means is that a component of the infrastructure, namely the PCA Manager, is deployed on all devices participating in the system setup. This implies that users and collaborating agents are able to participate in practice-centred cross-boundary decision support by sharing information (as peers in a distributed smart workspace) in the context of a user’s work setting and problem-solving requirements. Essentially, the device that generates a percept within the user’s work setting or that accepts user query in any cycle of cross-boundary decision support acts as the de facto server or “super-peer” [6]. Whereas a query-driven cycle, i.e. a cycle of decision support initiated by user query, follows the traditional request-response model, a percept-driven cycle follows an action-notify type of interaction, in which “perceived” changes in user work setting invoke appropriate handlers that ultimately result in alert messages for cross-boundary decision support.

The main components of the PCA Manager include the Work Practice Modelling component and the ContextMorph engine [4, 15]. The Work Practice Modelling engine organises a work description (WPD) into three constituent parts, namely domain factors, stereotypes about how the work is performed in the given locality, and perceived entities and their relationships within the work setting. The ContextMorph engine adapts a remote suggestion to user work setting using three low-level processes: similarity analysis, suggestion augmentation and morphing. The client end comprises two components: the cross-boundary collaboration layer and the UI layer. The collaboration layer handles interactions between user and collaborating agents across work boundaries, whereas the UI layer handles the visual displays of awareness and decision support information. The implementation processes and usage of the components will be described in details in the next section. The backend layer of CaDHealth comprises four distinct memory types (see Fig. 4).

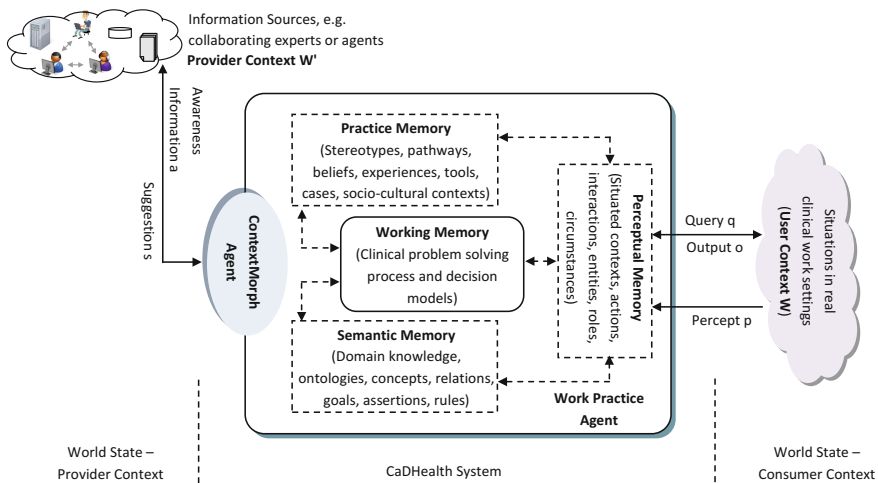


Fig. 4. CaDHealth architecture illustrating the system memory

Our aim is to represent memory in CaDHealth as a “boundary artifact” that holds the state of a work process [21, p. 63]. As shown in Fig. 4, the system memory includes *the working memory*, *the semantic memory*, *the practice memory*, and *the perceptual memory*. The working memory is the part of the system holds information that is actively in clinical problem-solving, reasoning and decision making, and makes it available to other parts of the system. An example of such information is a decision model. The semantic memory stores domain and concept-based knowledge related to a clinical work process. The third category of memory in CaDHealth, the practice memory, stores traces of an organisation’s past activities, practices (i.e. work approaches) employed to perform the activities, and experiences. The idea of practice memory resonates with Ackerman and Halverson’s [21] notion of organisational memory. However, practice memory is more holistic in its approach since it embodies the ontological, stereotyped and situated factors of an organisational work process and patterns. Examples of information stored in the practice memory include practice models and their instances, e.g. work stereotypes, pathways, clinical cases, work experience information. The perceptual memory stores information about entities perceived in a work environment, their categories and relationships between them. As they noted, the representational state of a work context, as stored in the working memory, in moving across organisational and regional boundaries, must necessarily lose some of its context. As a result, the ContextMorph agent, performs the decontextualisation and recontextualisation processes required to ensure that sufficient details about user work context is passed across boundary to the remote agent and that the suggestion provided by the remote agent is adapted to suit user work context and problem requirements.

3.2 Implementation: A Scenario Example of Using CaDHealth

To demonstrate how CaDHealth is used to provide PCA of a clinical work situation to a remote collaborating agent and offer enriched information to support user decision-making based on the agent’s suggestion, we describe an example of the chain of events that are involved in the interactive use of CaDHealth to provide cross-aware support in clinical decision making. The activity diagram in Fig. 5 shows what occurs when an unknown clinical work situation is presented to the system during an e-health decision support process, and illustrates the core functions of CaDHealth, which are described in greater details in [15].

The starting point for a cycle of decision support in CaDHealth occurs when a user sends a query seeking assistance from a remote agent or when the system automatically perceives a work environment and detects changes in entities’ conditions that require further information. This is handled by the functions: `getPercept()` and `sendQuery()` respectively. The next step is to update the work practice model based on the new information contained in the user query or system generated percept. Recall that at design time, the functions: `getOntologicalPractice()`, `getStereotype()` and `getSituatingPractice()` were used to construct the initial *WPD*. During update, the functions are re-run based on the new information. On the basis of the new *WPD*, the system generates work practice awareness information by executing the function: `generatePracticeInfo()`, which is sent to a remote collaborating agent as practice display.

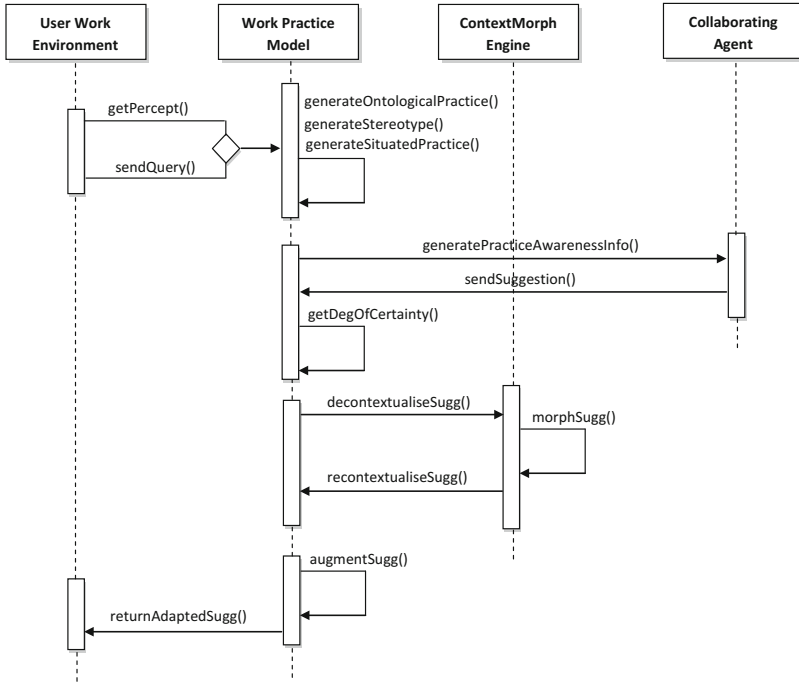


Fig. 5. Activity diagram showing a decision cycle in CaDHealth. ContextMorph Engine is described in [15]

When an agent sends a suggestion, CaDHealth executes the function `sendSuggestion()`, and consequently, determines the suggestion’s degree of certainty by executing the function: `getDegOfCertainty()`. In order to adapt the remote suggestion to user’s work practice, the ContextMorph engine executes three functions: `decontextualiseSugg()`, `morphSugg()` and `recontextualiseSugg()`. Finally the work practice model executes the `augmentSugg()` function, and, by executing the function `returnAdaptedSugg()`, returns the enriched suggestion to the user for decision support (Fig. 5). The simulation of these functions was implemented in Brahms [16], and described further in [15].

Note that CaDHealth is designed to be used by medical professionals, not patients, and in the context of chronic disease management as opposed to emergency situations – with the goal of enabling clinicians to obtain from across their organisational and regional boundaries peer opinion that would enable them to provide for their patients the best possible care tailored to their work practices and patients’ needs. Hence, CaDHealth includes a log in interface as shown in Fig. 6. After a successful login, the user is taken to the main window (Fig. 6). This window contains options for setting up a new case and viewing CaDHealth case analysis summary, for viewing ranked list of suggestions for decision support, and for chatting with remote experts. CaDHealth is able to generate practice-aware decision information and match information provided during new case setup to ontological, stereotyped and situated information obtained from the system database.

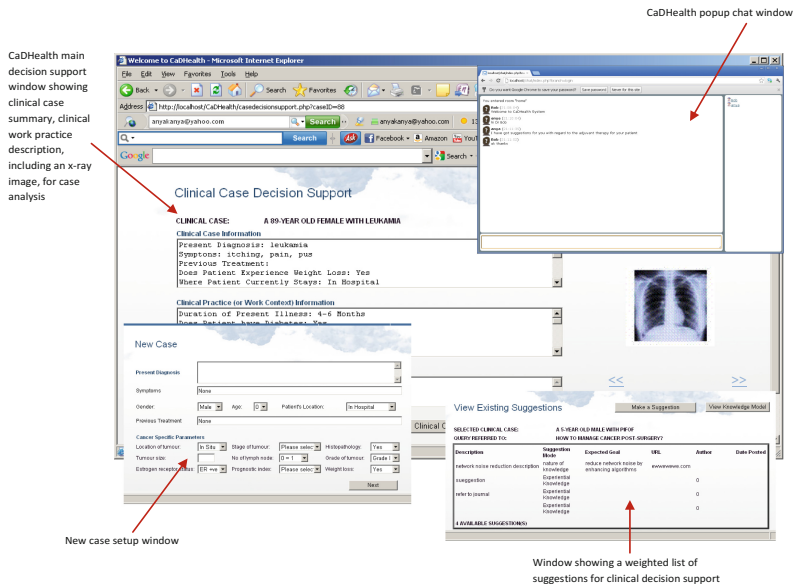


Fig. 6. CaDHealth web application user interface

During chat, CaDHealth is able to generate a suggestion from a remote expert’s chat message by relying on the user’s annotation of the chat message as well as on information obtained from work practice model for the particular clinical case. System evaluation is presented in [22], and identifies primarily integration of awareness information into an organisation’s clinical pathways as a major issue in designing support for awareness in e-health systems [23, 24].

4 Conclusion

This paper presents the design of an awareness environment for cross-boundary clinical decision support in e-health that takes account of the concept of work practice as a design requirement of e-health decision support systems. A key benefit of the proposed modelling approach is that it offers newer and more broad-based perspectives for gaining deeper understanding of people’s interaction with technologies in real-world problem-solving situations. By integrating context into a combined model of human activity system and situation awareness, the proposed model provides novel ways of highlighting the nuances in real-world problem-solving across work settings and enables sense making of work activities beyond what are offered by existing workflow-based models, and thus results potentially in more effective cross-boundary decision support for e-health. It allows for the use of logical reasoning (domain rules and ontologies), probability factors (for stereotyping) and situatedness (case-based reasoning) to enable reasoning about work contexts. However, the system does not sufficiently provide support for intelligent techniques that allow “learning on the job”,

and does not take account of the influence of such factors as body language, emotions and expressions on decisions; these issues will form the central focus of our future work.

References

1. Heath, C., Svensson, M.S., Hindmarsh, J., Luff, P., vom Lehn, D.: Configuring awareness. *CSCW* **11**(3–4), 317–347 (2002)
2. Vyas, D.: Designing for awareness: an experience-focused HCI perspective. Ph.D. Dissertation, University of Twente, The Netherlands (2011)
3. Riemer, K., Haines, R.: Pools and streams: a theory of dynamic, practice-based awareness creation in mediated-communication. In: JAIS Theory Development Workshop, Sprouts: Working Papers on Information Systems, vol. 8, no 12, (2008). http://aisel.aisnet.org/sprouts_all/217/. Accessed 15 Nov 2017
4. Anya, O., Tawfik, H., Nagar, A., Amin, S., Shaalan, K.: Context-aware knowledge modelling for decision support in e-Health. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN), the IEEE Computational Intelligence, Barcelona, Spain, pp. 1–7 (2010)
5. Bardram, J.E., Hansen, T.R.: Context-based workplace awareness. *J. CSCW* **19**(2), 105–138 (2010)
6. Bardram, J., Gueddana, S., Houben, S., Nielsen, S.: ReticularSpaces: activity-based computing support for physically distributed and collaborative smart spaces. In: CHI 2012, pp. 2845–2854 (2012)
7. Dourish, P., Bellotti, V.: Awareness and coordination in shared workspaces. In: ACM Conference on CSCW, Toronto, Canada, pp. 107–114 (1992)
8. Brézillon, P.: Context-based management of practices. In: International Workshop on Modeling and Reasoning in Context, Karlsruhe, Germany. LNCS. Springer, Heidelberg, 26–30 September 2011. <http://events.idi.ntnu.no/mrc2011/papers.php>
9. Gutwin, C.: Workspace awareness in real-time distributed groupware. Ph.D. thesis, University of Calgary, Canada (1997)
10. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **37**(1), 32–64 (1995)
11. Schmidt, K.: The problem with ‘Awareness’: introductory remarks on ‘Awareness in CSCW’. *J. CSCW* **11**(3–4), 285–298 (2002)
12. You, Y.: A survey for the study of awareness in co-operative systems. In: Conference on Information Systems Research in Scandinavia (IRIS 23), Uddevalla, Sweden (2000)
13. Clancey, W.J.: Observation of work practices in natural settings. In: Ericsson, A., Charness, N., Feltovich, P., Hoffman, R. (eds.) *Cambridge Handbook on Expertise and Expert Performance*, pp. 127–145. Cambridge University Press, New York (2006)
14. Gabbay, J., le May, A.: *Practice-Based Evidence for Healthcare*. Routledge, New York (2011)
15. Anya, O.: Practice-centred e-health system design for cross-boundary clinical decision support. Ph.D. thesis, University of Liverpool, UK (2012)
16. Sierhuis, M., Clancey, W.J., van Hoof, R.J.J.: Brahms an agent-oriented language for work practice simulation and multi-agent systems development. In: El Fallah Seghrouchni, A., Dix, J., Dastani, M., Bordini, R.H. (eds.) *Multi-Agent Programming*, pp. 73–117. Springer, Boston, MA (2009). https://doi.org/10.1007/978-0-387-89299-3_3

17. Engeström, Y.: Learning by expanding: an activity-theoretical approach to developmental work research. Orienta-konsultit, Helsinki (1987)
18. Novak, J.: Helping knowledge cross boundaries: using knowledge visualization to support cross-community sensemaking. In: HICSS, Waikoloa, HI, USA (2007)
19. Huang, Z., Chen, H., Guo, F., Xu, J.J., Wu, S., Chen, W.: Expertise visualization: an implementation and study based on cognitive fit theory. *Decis. Support Syst.* **42**(3), 1539–1557 (2006)
20. IOM, Committee on Quality of Health Care in America: Crossing the Quality Chasm: A New Health System for the 21st Century. National Academy Press, Washington, D.C. (2001)
21. Ackerman, M.S., Halverson, C.A.: Re-examining organizational memory. *Commun. ACM* **43**(1), 58–64 (2000)
22. Tawfik, H., Anya, O.: Evaluating practice-centered awareness in cross-boundary telehealth decision support systems. *Telemat. Inform.* **32**, 486–503 (2015)
23. Tawfik, H., Anya, O., Nagar, A.: Understanding clinical work practices for cross-boundary decision support in e-Health. *IEEE Trans. Inf. Technol. Biomed.* **16**(4), 530–541 (2012)
24. Daniyal, A., Abidi, S.R., Abidi, S.S.R.: Computerizing clinical pathways: ontology-based modeling and execution. In: MIE, pp. 643–647 (2009)
25. Anya, O., Tawfik, H., Nagar, A., Amin, S.: Context-aware decision support in knowledge-intensive collaborative e-Work. In: Proceedings of the International Conference on Computational Science (ICCS), pp. 2281–2290 (2010)



Correction to: Textile Sensor Platform (TSP) - Development of a Textile Real-Time Electrocardiogram

Thomas Walzer, Christian Thies, Klaus Meier,
and Natividad Martínez Madrid

Correction to:
**Chapter “Textile Sensor Platform (TSP) - Development
of a Textile Real-Time Electrocardiogram”**
**in: I. Rojas and F. Ortuño (Eds.): *Bioinformatics
and Biomedical Engineering*, LNBI 10814,**
https://doi.org/10.1007/978-3-319-78759-6_33

The acknowledgement section of this paper was not given originally. It was added in the paper to thank some students who contributed to the work but were not authors.

The updated version of this chapter can be found at
https://doi.org/10.1007/978-3-319-78759-6_33

© Springer International Publishing AG, part of Springer Nature 2019
I. Rojas and F. Ortuño (Eds.): IWBBIO 2018, LNBI 10814, p. C1, 2018.
https://doi.org/10.1007/978-3-319-78759-6_42

Author Index

- Abou Tabl, Ashraf I-166
Abouelhoda, Mohamed I-405
Acosta, Antonio I-96
Adebiyi, Marion O. I-290
Adetiba, Emmanuel I-290, II-266
Ahmad, Jamil I-543
Ahmaidi, Said II-59, II-128
Akanle, Matthew B. I-290, II-266
Akinrinmade, Adekunle II-266
Al Meer, Hossam I-71
Alberca Díaz-Plaza, Ana II-135
Albuquerque Nogueira, Romildo I-57
Al-Harbi, Najla I-197
Ali, Amjad I-543
Aljafar, Hussain I-405
Alkhateeb, Abedalrhman I-166, I-343
Alnakhli, Yasser I-405
Alpala-Alpala, Luis Omar II-231
Alpar, Orcan II-243, II-255
Alsbeih, Ghazi I-197
Alshammari, Sultanah M. II-430
Álvarez, Óscar I-211
Alvarez-Machancoses, Óscar II-15, II-24, II-33
Alvarez-Uribe, Karla C. II-289
Alves, P. I-329
Alwan, Abir II-69, II-128
Aminian, Kamiar II-75
Angles, Renzo I-235
Anjum, Ashiq I-405
Antal, Peter I-41
Antelis, Javier Mauricio I-129
Antsiperov, Viacheslav II-383
Anya, Obinna II-456
Arciniegas-Mejia, Andrés F. I-96
Arenas, Mauricio I-235
Arrais, Joel P. I-221
Arroyo, Macarena I-373
Årsand, Eirik II-443
Arzalluz-Luque, Angeles I-364
Ayoub, Marie-Louise II-59

Bachratá, Katarína I-279
Bachratý, Hynek I-279

Badejo, Joke A. I-290, II-266
Bárta, Antonín I-3, I-107, I-139
Basile, Teresa M. A. I-302
Bastidas Torres, D. I-26
Batista, Cassio I-415
Bautista, Rocío I-373
Becerra, Miguel Alberto I-26, II-289
Becerra-Botero, Miguel A. II-231
Bekkhzhayeva, Dinara I-139
Bekkozhayeva, Dinara I-3
Bekkozhayeva, Dinara I-107
Bellotti, Roberto I-302
Benjelloun, Mohammed II-198
Bensmail, Halima I-71
Bereksi Reguig, Fethi II-311
Berro, Abdel-Jalil II-59, II-69, II-128
Betrouni, Nacim II-409
Beuscart, Régis II-421
Blanco Valencia, X. I-26
Blazek, Pavel I-352
Bobrowski, Leon I-153
Bodin, Oleg N. II-325
Bohiniková, Alžbeta I-259
Bohoyo, Pablo de Miguel II-135
Bottigli, Ubaldo I-302
Boudet, Samuel II-421
Boyano, Maria Dolores I-511
Bozhynov, Vladyslav I-3, I-107, I-139
Bringas, Carlos I-511
Brunscics, Bence I-41

Canakoglu, Arif I-270
Caon, Maurizio II-75
Cárdenas-García, Maura I-83
Castellanos-Domínguez, Cesar Germán II-231
Castro-Ospina, Andrés Eduardo I-26, I-96
Cekas, Elingas II-49
Cernea, Ana I-211, II-15, II-24, II-33
Cerqueira, Fabio R. I-383
Chen, Bolin I-247
Chemikov, Anton I. II-325
Chovanec, Michal I-279
Cimrák, Ivan I-259

- Císar, Petr I-139
 Claros, M. Gonzalo I-373
 Coelho, Edgar D. I-221
 Coimbra, P. I-329
 Conti, Massimo II-347
 Cortet, Bernard II-59
 Cruz-Cruz, Lilian Dayana II-231
 Cyran, Norbert I-139
- da Silva, José Radamés Ferreira I-57
 da Silveira, Carlos H. I-383
 Dandekar, Thomas I-395
 de l'Aulnoit, Agathe Houzé II-421
 de l'Aulnoit, Denis Houzé II-421
 de Melo-Minardi, Raquel C. I-383
 De Paolis, Lucio Tommaso II-118
 de Andrés-Galiana, Enrique J. II-15, II-24,
 II-33
 Demailly, Romain II-421
 Dentamaro, Rosalba I-302
 Devailly, Guillaume I-364
 Dey, Kashi Nath II-397
 Díaz, Gloria M. II-106
 Díaz-Del-Pino, Sergio I-177
 Díaz-del-Pino, Sergio I-450
 Didonna, Vittorio I-302
 Dobruch-Sobczak, Katarzyna II-186
 Drisis, Stylianos II-198
- El Hage, Rawad II-59, II-69
 El Houry, César II-59, II-69
 El Houry, Georges II-59
 El-Kalioby, Mohamed I-405
 ElMaraghy, Waguih I-166
 Esseiva, Julien II-75
 Estella, Francisco II-176
- Fanizzi, Annarita I-302
 Faquih, Tariq I-405
 Fassio, Alexandre V. I-383
 Fathi, Madjid I-531
 Fausto, Alfonso I-302
 Fayad, Ibrahim II-69
 Feng, Xuan I-461
 Fernández-Martínez, Juan Luis I-211, II-15,
 II-24, II-33
 Fernández-Muñiz, Zulima I-211, II-15,
 II-24, II-33
 Fernández-Ovies, Francisco Javier II-15,
 II-24, II-33
- Ferreira da Rocha, Adson II-95
 Fethi, Bereksi-Reguig II-301
 Fostier, Jan I-439
 Fotouhi, Ali I-429
 Frenn, Fabienne II-128
- Gaiduk, Maksym II-347, II-371
 García, Pedro I-96
 García-Rojo, Marcial II-276
 Gardeazabal, Jesus I-511
 Garzon, Max H. I-486
 Geng, Yu I-473
 Genin, Michael II-421
 Gezsi, Andras I-41
 Ghosh, Anupam II-397
 Gil, M. H. I-329
 Golcuk, Guray I-270
 González, Jesús II-276
 González, Ramón E. R. I-57
 González-Castaño, Catalina II-231
 González-Pérez, Pedro Pablo I-83
 Guan, Xin II-3
 Gudnason, Kristinn I-329
 Guiomar, A. J. I-329
 Guta, Gabor I-41
- Hadj Henni, Abderraouf II-409
 Hage, Rawad El II-128
 Hamzeh, Osama I-343
 Hargaš, Libor II-163
 Hartvigsen, Gunnar II-443
 Helsper, Julien I-531
 Hernández, Luis Guillermo I-129
 Holthausen, Ricardo I-177, I-450
 Hu, Huan I-461
- Imbajoa-Ruiz, David Esteban I-96
 Isik, Zerrin I-520
- Jablončík, František II-163
 Janusas, Giedrius II-49
 Jonsdottir, Fjola I-329
 Joshi, Anagha I-364
 Judia, Sara Bin I-197
 Jung, Matthias I-531
- Kajánek, František I-279
 Karwat, Piotr II-186
 Khaled, Omar Abou II-75
 Kirimat, Ayca II-212, II-221

- Klimonda, Ziemowit II-186
 Kloczkowski, Andrzej I-211
 Koniar, Dušan II-163
 Korbicz, Józef II-151
 Kovalčíková, Kristína I-259
 Kowal, Marek II-151
 Kramm, Mikhail N. II-325
 Krejcar, Ondrej I-352, II-212, II-221, II-243, II-255
 Kuca, Kamil I-352
 Külekci, M. Oğuzhan I-429
 Kuonen, Pierre I-395
 Kupriyanova, Yana A. II-325
- La Forgia, Daniele I-302
 Lachkar, Abdelmonaime I-314
 Larhman, Mohamed Amine II-198
 Larrosa, Rafael I-373
 Leber, Isabel II-335
 Legarreta, Leire I-511
 Li, Yang I-473
 Litniewski, Jerzy II-186
 Liu, Li II-3
 Loncová, Zuzana II-163
 López, Vanessa I-15
 Lopez-Chamorro, Fabián M. I-96
 López-Delis, Alberto II-95, II-106
 López-Rodríguez, Carmen María I-373
 Losurdo, Liliana I-302
 Lozano, Beatriz II-176
 Luo, Ping I-247
- Maalouf, Ghassan II-69, II-128
 Maarouf, Haitham I-15
 Madrid, Natividad Martínez II-335, II-347, II-359
 Mahmoudi, Saïd II-198
 Majid, Salma I-197
 Majidi, Mina I-429
 Malaina, Iker I-511
 Mall, Raghvendra I-71
 Malyshev, Andrew I-186
 Mansurov, Gennady II-383
 Martínez de la Fuente, Ildelfonso I-511
 Martínez, Diego I-15
 Martínez, Efraín I-129
 Martínez, Luis I-511
 Massafra, Raffaella I-302
 Matta, Joseph II-128
- Matthews, Victor O. I-290
 Mazza Guimaraes, Isabelle I-259
 Megrez, Nasreddine I-71
 Meier, Klaus II-359
 Mikler, Armin R. II-430
 Miranda, Fábio I-415
 Monaco, Alfonso I-302
 Monczak, Roman II-151
 Morais, Jefferson I-415
 Mora-Jiménez, Inmaculada II-135
 Moschetta, Marco I-302
 Mugellini, Elena II-75
 Muñoz-Minjares, Jorge II-85
- Nabil, Dib II-301
 Nadia, Ouacif II-301
 Nasiri, Sara I-531
 Nasr, Riad II-69
 Neto, Nelson I-415
 Ngom, Alioune I-166
- Oliveira, José Luís I-221
 Oller, Josep M. I-501
 Olugbara, Oludayo O. I-290
 Orcioni, Simone II-347
 Ortega, Juan Antonio II-371
 Ortega, Julio II-276
- Palevicius, Arvydas II-49
 Pasquier, David II-409
 Patel, YatinkumarRajeshbhai II-49
 Peluffo-Ordóñez, Diego Hernán I-26, I-96, II-231, II-289
 Penzel, Thomas II-371
 Perez-Chimal, R. J. II-85
 Pérez-Wohlfeil, Esteban I-177, I-450
 Peyrodie, Laurent II-421
 Pham, Duy T. I-486
 Piñeros Rodríguez, C. I-26
 Pinti, Antonio II-59, II-69, II-128, II-421
 Piotrkowska-Wróblewska, Hanna II-186
 Polanska, Joanna I-197
 Popescu, Ondina I-302
 Popova, Tatiana II-85
- Rais, Mohammed I-314
 Ramos, Rommel I-415
 Ramos-López, Javier II-135
 Redouane, Benali II-301

- Reverter, Ferran I-501
 Rizkallah, Maroun II-128
 Rodríguez-Brazzarola, Pablo I-177, I-450
 Rojas, Fernando II-176
 Rojas, Ignacio II-176
 Romanelli, João P. R. I-383
 Rønningen, Ida Charlotte II-443
 Rosero-Montalvo, Paul D. I-96
 Roy, Sukriti II-397
 Rubio-Sánchez, Manuel II-135
 Rueda, Luis I-166, I-343
 Ruíz-Olaya, Andrés F. II-95, II-106
- Saddik, Hayman II-69, II-128
 Sadovsky, Michael I-186
 Saeed, Muhammad Tariq I-543
 Saha, Sujay II-397
 Saiz, Antonio II-176
 Salazar-Castro, Jose Alejandro II-231
 Saligan, Leorey II-15, II-24, II-33
 Sánchez, Alberto II-135
 Sanchez-Morillo, Daniel II-276
 Santamarta, Elena II-176
 Santana, Charles A. I-383
 Schiro, Jessica II-421
 Sedjelmaci, Ibticeme II-311
 Seepold, Ralf II-325, II-347, II-371
 Seijo, Fernando II-176
 Senashova, Maria I-186
 Shah, Zeeshan Ali I-405
 Shmaliy, Yuriy S. II-85
 Shokrof, Moustafa I-405
 Sigurdsson, Sven I-329
 Silva, Artur I-415
 Silveira, Sabrina de A. I-383
 Skobel, Marcin II-151
 Slavík, Martin I-259, I-279
 Smiešková, Monika I-279
 Sobrido, Maria J. I-15
 Soguero-Ruiz, Cristina II-135
 Sonis, Stephen T. II-15, II-24, II-33
 Souček, Pavel I-3, I-107, I-139
- Suarez, Esther II-176
 Subhani, Shazia I-405
- Taboada, Maria I-15
 Tahar, Omari II-301
 Taiwo, Tunmike B. I-290
 Tamborra, Pasquale I-302
 Tangaro, Sabina I-302
 Tawfik, Hissam II-456
 Thies, Christian II-359
 Tian, Li-Ping I-247
 Tobiasz, Joanna I-197
 Tolan, Ertan I-520
 Trelles, Oswaldo I-177, I-450
 Tuncel, Mustafa Anil I-270
- Ullah, Ehsan I-71
 Umaquina-Criollo, Ana Cristina II-231
 Urban, Jan I-118
 Urbanová, Pavla I-3, I-107, I-139
- Vegas, Esteban I-501
 Vieira, A. P. I-329
 Volák, Jozef II-163
 Vunderl, Bruno II-371
- Walzer, Thomas II-359
 Wang, Jiayin I-461, I-473
 Watelain, Eric II-69
 Wolf, Beat I-395
 Wu, Fang-Xiang I-247
- Xiao, Qianghua I-247
 Xiao, Xiao I-473
- Zabielski, Paweł I-153
 Zakhem, Eddy II-59
 Železný, Miloš I-139
 Zhang, Xuanping I-461, I-473
 Zhao, Zhongmeng I-461, I-473
 Zheng, Tian I-473
 Zhikhareva, Galina V. II-325
 Zhuravleva, Natalija A. II-325