

Yen-Wei Chen  
Satoshi Tanaka  
Robert J. Howlett  
Lakhmi C. Jain *Editors*



# Innovation in Medicine and Healthcare 2016



# **Smart Innovation, Systems and Technologies**

**Volume 60**

## **Series editors**

Robert James Howlett, KES International, Shoreham-by-sea, UK  
e-mail: rjhowlett@kesinternational.org

Lakhmi C. Jain, University of Canberra, Canberra, Australia;  
Bournemouth University, UK;  
KES International, UK  
e-mails: jainlc2002@yahoo.co.uk; Lakhmi.Jain@canberra.edu.au

### *About this Series*

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

More information about this series at <http://www.springer.com/series/8767>

Yen-Wei Chen · Satoshi Tanaka  
Robert J. Howlett · Lakhmi C. Jain  
Editors

# Innovation in Medicine and Healthcare 2016



Springer

*Editors*

Yen-Wei Chen  
College of Information Science  
and Engineering  
Ritsumeikan University  
Kusatsu, Shiga  
Japan

Satoshi Tanaka  
Department of Media Technology  
Ritsumeikan University  
Kusatsu, Shiga  
Japan

Robert J. Howlett  
KES International  
Shoreham-by-sea  
UK

Lakhmi C. Jain  
University of Canberra  
Canberra  
Australia

and

Bournemouth University  
Poole  
UK

and

KES International  
Shoreham-by-sea  
UK

ISSN 2190-3018

ISSN 2190-3026 (electronic)

Smart Innovation, Systems and Technologies

ISBN 978-3-319-39686-6

ISBN 978-3-319-39687-3 (eBook)

DOI 10.1007/978-3-319-39687-3

Library of Congress Control Number: 2016941079

© Springer International Publishing Switzerland 2016

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

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

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

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

# Preface

The 4th KES International Conference on Innovation in Medicine and Healthcare (InMed-16) was held during 15–17 June 2016 in Tenerife, Spain, organized by KES International.

The InMed-16 is the 4th edition of the InMed series of conferences. The 1st, the 2nd and the 3rd InMed Conferences were held in Italy, Spain and Japan, respectively. The conference focuses on major trends and innovations in modern intelligent systems applied to medicine, surgery, healthcare and the issues of an ageing population. The purpose of the conference is to exchange the new ideas, new technologies and current research results in these research fields.

We received submissions from more than ten countries. All submissions were carefully reviewed by at least two reviewers of the International Programme Committee. Finally 32 papers were accepted to be presented in these proceedings. The major areas covered at the conference and presented in these proceedings include: (1) Innovative Technology in Mental Healthcare; (2) Intelligent Decision Support Technologies and Systems in Healthcare; (3) Biomedical Engineering, Trends, Research and Technologies; (4) Advances in Data and Knowledge Management for Healthcare; (5) Advanced ICT for Medical and Healthcare; (6) Healthcare Support System; and (7) Smart Medical and Healthcare System.

We would like to thank Dr. Kyoko Hasegawa and Ms. Yuka Sato of Ritsumeikan University for their valuable editing assistance for this book. We are also grateful to the authors and reviewers for their contributions.

June 2016

Yen-Wei Chen  
Satoshi Tanaka  
Robert J. Howlett  
Lakhmi C. Jain

# **InMed 2016 Organization**

## **General Co-Chairs**

Lakhmi C. Jain, University of Canberra, Canberra, Australia and Bournemouth University, UK  
Robert J. Howlett, KES International, Shoreham-by-sea, UK

## **Programme Co-Chairs**

Yen-Wei Chen, Ritsumeikan University, Japan  
Satoshi Tanaka, Ritsumeikan University, Japan

## **Co-Chairs of Workshop on Smart Medical and Healthcare Systems**

Elder Sanchez, Vicomtech-IK4, Spain  
Ivan Macia, Vicomtech-IK4, Spain

## **International Programme Committee Members**

Ahmad Taher Azar, Benha University, Egypt  
Vitoantonio Bevilacqua, DEI, Politecnico di Bari, Italy  
Isabelle Bichindaritz, State University of New York at Oswego, USA  
Giosuè Lo Bosco, University of Palermo, Italy  
Sheryl Brahnam, Computer Information Systems Missouri State University, USA  
Christopher Buckingham, Aston University, UK  
Antonio Fernández-Caballero, Universidad de Castilla-La Mancha, Spain  
Nashwa El-Bendary, Arab Academy for Science, Technology, and Maritime Transport, Giza, Egypt

Massimo Esposito, National Research Council of Italy (ICAR-CNR), Italy  
Jesús M. Doña Fernández, Andalusian Health Service, Spain and National Distance Education University (UNED), Spain  
María del Rosario Baltazar Flores, Instituto Tecnológico de León, México  
José Manuel Matos Ribeiro da Fonseca, NOVA University of Lisbon, Portugal  
Arnulfo Alanis Garza, Instituto Tecnológico de Tijuana, Mexico  
Arfan Ghani, University of Bolton, Greater Manchester, UK  
Manuel Graña, Universidad del País Vasco, Spain  
Ayako Hashizume, Tokyo Metropolitan University, Japan  
Ioannis Hatzilygeroudis, University of Patras, Greece  
Yasushi Hirano, Yamaguchi University, Japan  
Ajita Ichalkaranje, Physiotherapist, Seven Steps Physiotherapy Clinic, Adelaide, South Australia  
Sandhya Jain, Medical Practitioner, Adelaide, South Australia  
Vijay Mago, Lakehead University, Canada  
Andrea Matta, Shanghai Jiao Tong University, China  
Rashid Mehmood, King Abdul Aziz University, Saudi Arabia  
José Miguel Latorre, Universidad de Castilla-La Mancha, Centro de Investigación en Criminología  
Aniello Minutolo, Institute for High Performance Computing and Networking—National Research Council of Italy (ICAR-CNR)  
Louise Moody, Coventry School of Art and Design, Coventry University, UK  
Antonio Moreno, Universitat Rovira i Virgili, Spain  
Marek R. Ogiela, AGH University of Science and Technology, Krakow, Poland  
Andrés Ortiz, Universidad de Málaga, Spain  
Elpiniki Papageorgiou, Technological Educational Institute of Lamia, Greece  
José Manuel Pastor, Universidad de Castilla-La Mancha, Instituto de Tecnologías Audiovisuales  
Elena Hernández Pereira, University of A Coruña, Spain  
Luigi Portinale, University of Piemonte Orientale, Italy  
Dorin Popescu, University of Craiova, Romania  
Marco Pota, Institute for High Performance Computing and Networking—National Research Council of Italy (ICAR-CNR)  
Ana Respício, Universidade de Lisboa, Portugal  
Roberto Rodríguez-Jiménez, Psychiatry, Hospital  
Joel J.P.C. Rodrigues, Instituto de Telecomunicações, University of Beira Interior, Portugal  
John Ronczka, SCOTTYNCC Independent Research Scientist, Australia  
Yves Rybarczyk, Nova University of Lisbon FCT/UNL, Portugal  
Marina V. Sokolova, Southwest State University, Russia  
Yoshiyuki Yabuuchi, Shimonoseki City University, Japan  
Hiroyuki Yoshida, Harvard Medical School / Massachusetts General Hospital, USA  
Lenin G. Lemus-Zúñiga, Universitat Politècnica de València. España

## **Organization and Management**

KES International ([www.kesinternational.org](http://www.kesinternational.org)) in partnership with the Institute of Knowledge Transfer ([www.ikt.org.uk](http://www.ikt.org.uk)).

# Contents

## Part I Innovative Technology in Mental Healthcare

<b>Non-lineal EEG Modelling by Using Quadratic Entropy for Arousal Level Classification . . . . .</b>	<b>3</b>
Arturo Martínez-Rodrigo, Raúl Alcaraz, Beatriz García-Martínez, Roberto Zangróniz and Antonio Fernández-Caballero	
<b>Emotional Induction Through Films: A Model for the Regulation of Emotions . . . . .</b>	<b>15</b>
Luz Fernández-Aguilar, José Miguel Latorre, Laura Ros, Juan Pedro Serrano, Jorge Ricarte, Arturo Martínez-Rodrigo, Roberto Zangróniz, José Manuel Pastor, María T. López and Antonio Fernández-Caballero	
<b>Application of the Lognormal Model to the Vocal Tract Movement to Detect Neurological Diseases in Voice . . . . .</b>	<b>25</b>
Cristina Carmona-Duarte, Réjean Plamondon, Pedro Gómez-Vilda, Miguel A. Ferrer, Jesús B. Alonso and Ana Rita M. Londral	
<b>Exploring and Comparing Machine Learning Approaches for Predicting Mood Over Time . . . . .</b>	<b>37</b>
Ward van Breda, Johnno Pastor, Mark Hoogendoorn, Jeroen Ruwaard, Joost Asselbergs and Heleen Riper	
<b>Cross-Cultural Telepsychiatry: An Innovative Approach to Assess and Treat Ethnic Minorities with Limited Language Proficiency . . . . .</b>	<b>49</b>
Davor Mucic	

## Part II Biomedical Engineering, Trends, Research and Technologies

<b>A Comparison of Performance of Sleep Spindle Classification Methods Using Wavelets . . . . .</b>	<b>61</b>
Elena Hernandez-Pereira, Isaac Fernandez-Varela and Vicente Moret-Bonillo	

<b>Evaluation of Head-Mounted Displays for Macular Degeneration: A Pilot Study . . . . .</b>	71
Howard Moshtael, Lanxing Fu, Ian Underwood and Baljean Dhillon	
<b>Method of Infrared Thermography for Earlier Diagnostics of Gastric Colorectal and Cervical Cancer . . . . .</b>	83
B. Dekel, A. Zilberman, N. Blaunstein, Y. Cohen, M.B. Sergeev, L.L. Varlamova and G.S. Polishchuk	
<b>Tijuana's Sustainability for Healthcare Measurement Using Fuzzy Systems . . . . .</b>	93
Bogart Yail Márquez, Arnulfo Alamis, Jose Sergio Magdaleno-Palencia, Karina Romo, Felma González and Sergio Mendez-Mota	
<b>Intelligent System for Learning of Comfort Preferences to Help People with Mobility Limitations . . . . .</b>	99
Sandra López, Rosario Baltazar, Lenin Lemus Zuñiga, Miguel Ángel Casillas, Arnulfo Alanis, Miguel A. Mateo Pla, Víctor Zamudio and Guillermo Méndez	
<b>Design of a Middleware and Optimization Algorithms for Light Comfort in an Intelligent Environment . . . . .</b>	111
Teresa Barrón Llamas, Rosario Baltazar, Miguel A. Casillas, Lenin Lemus, Arnulfo Alanis and Víctor Zamudio	
<b>Autism Disorder Neurological Treatment Support Through the Use of Information Technology . . . . .</b>	123
Esperanza Manrique Rojas, Margarita Ramírez Ramírez, Hilda Beatriz Ramírez Moreno, Maricela Sevilla Caro, Arnulfo Alanís Garza and José Sergio Magdaleno Palencia	
<b>Information Technologies in the Area of Health and Use of Mobile Technologies in the Area of Health in Tijuana, Baja California . . . . .</b>	129
Margarita Ramírez Ramírez, Esperanza Manrique Rojas, Nora del Carmen Osuna Millán, María del Consuelo Salgado Soto, José Sergio Magdaleno Palencia and Bogart Yail Márquez Lobato	
<b>Part III Advances in Data and Knowledge Management for Healthcare</b>	
<b>Automatic Quantification of the Extracellular Matrix Degradation Produced by Tumor Cells . . . . .</b>	137
Nadia Brancati, Giuseppe De Pietro, Maria Frucci, Chiara Amoruso, Daniela Corda and Alessia Varone	
<b>Quantitative EEG and Virtual Reality to Support Post-stroke Rehabilitation at Home . . . . .</b>	147
Alfonso Mastropietro, Sara Arlati, Simona Mrakic-Sposta, Luca Fontana, Cristina Franchin, Matteo Malosio, Simone Pittaccio, Cristina Gramigna, Franco Molteni, Marco Sacco and Giovanna Rizzo	

<b>Towards a Sustainable Solution for Collaborative Healthcare Research</b> . . . . .	159
Nikos Karacapilidis and George Potamias	
<b>An Ontology-Based Approach for Representing Medical Recommendations in mHealth Applications</b> . . . . .	171
Aniello Minutolo, Massimo Esposito and Giuseppe De Pietro	
<b>Semantic Cluster Labeling for Medical Relations</b> . . . . .	183
Anita Alicante, Anna Corazza, Francesco Isgrò and Stefano Silvestri	
<b>Part IV Advanced ICT for Medical and Healthcare</b>	
<b>A Robust Zero-Watermarking Algorithm for Encrypted Medical Images in the DWT-DFT Encrypted Domain</b> . . . . .	197
Jiangtao Dong and Jingbing Li	
<b>A Touchless Visualization System for Medical Volumes Based on Kinect Gesture Recognition</b> . . . . .	209
Ryoma Fujii, Tomoko Tateyama, Titinunt Kitrungrotsakul, Satoshi Tanaka and Yen-Wei Chen	
<b>Super-Resolution Technology for X-Ray Images and Its Application for Rheumatoid Arthritis Medical Examinations</b> . . . . .	217
Tomio Goto, Takuma Mori, Hidetoshi Kariya, Masato Shimizu, Masaru Sakurai and Koji Funahashi	
<b>Bayesian Model for Liver Tumor Enhancement</b> . . . . .	227
Yu Konno, Xian-Hua Han, Lanfen Lin, Hongjie Hu, Yitao Liu, Wenchao Zhu and Yen-Wei Chen	
<b>Fused Visualization with Non-uniform Transparent Surface for Volumetric Data Using Stochastic Point-Based Rendering</b> . . . . .	237
Kyoko Hasegawa, Yuta Fujimoto, Rui Xu, Tomoko Tateyama, Yen-Wei Chen and Satoshi Tanaka	
<b>Part V Healthcare Support System</b>	
<b>Automated Diagnosis of Parkinsonian Syndromes by Deep Sparse Filtering-Based Features</b> . . . . .	249
Andrés Ortiz, Francisco J. Martínez-Murcia, María J. García-Tarifa, Francisco Lozano, Juan M. Górriz and Javier Ramírez	
<b>Post-stroke Hand Rehabilitation Using a Wearable Robotic Glove</b> . . . . .	259
Dorin Popescu, Mircea Ivanescu, Razvan Popescu, Anca Petrisor, Livia-Carmen Popescu and Ana-Maria Bumbea	

<b>An mHealth Application for a Personalized Monitoring of One's Own Wellness: Design and Development . . . . .</b>	269
Manolo Forastiere, Giuseppe De Pietro and Giovanna Sannino	
<b>"White Coat" Effect Study as a Subclinical Target Organ Damage by Means of a Web Platform . . . . .</b>	279
J. Novo, A. Hermida, M. Ortega, N. Barreira, M.G. Penedo, J.E. López and C. Calvo	
<b>Part VI Smart Medical and Healthcare Systems 2016 Workshop</b>	
<b>GENESIS—Cloud-Based System for Next Generation Sequencing Analysis: A Proof of Concept . . . . .</b>	291
Maider Alberich, Arkaitz Artetxe, Eduardo Santamaría-Navarro, Alfons Nonell-Canals and Grégory Maclair	
<b>Review of Automatic Segmentation Methods of White Matter Lesions on MRI Data . . . . .</b>	301
Darya Chyzyk, Manuel Graña and Gerhard Ritter	
<b>Hygehos Ontology for Electronic Health Records . . . . .</b>	311
Naiara Muro, Eider Sanchez, Manuel Graña, Eduardo Carrasco, Fran Manzano, Jose María Susperregi, Agustín Agirre and Jesús Gómez	
<b>Views on Electronic Health Record . . . . .</b>	323
Manuel Graña and Oier Echaniz	
<b>Laparoscopic Video and Ultrasounds Image Processing in Minimally Invasive Pancreatic Surgeries . . . . .</b>	333
P. Sánchez-González, I. Oropesa, B. Rodríguez-Vila, M. Viana, A. Fernández-Pena, T. Arroyo, J.A. Sánchez-Margallo, J.L. Moyano, F.M. Sánchez-Margallo and E.J. Gómez	
<b>Erratum to: Hygehos Ontology for Electronic Health Records . . . . .</b>	E1
Naiara Muro, Eider Sanchez, Manuel Graña, Eduardo Carrasco, Fran Manzano, Jose María Susperregi, Agustín Agirre and Jesús Gómez	
<b>Author Index . . . . .</b>	345

**Part I**

**Innovative Technology in Mental  
Healthcare**

# Non-lineal EEG Modelling by Using Quadratic Entropy for Arousal Level Classification

Arturo Martínez-Rodrigo, Raúl Alcaraz, Beatriz García-Martínez,  
Roberto Zangróniz and Antonio Fernández-Caballero

**Abstract** Nowadays, assistive technologies together with ubiquitous and pervasive computing are emerging as main alternative to help ageing population. In this respect, an important number of works have been carried out to improve the quality of life in elderly from a physical point of view. However, less efforts have been made in monitoring the mental and emotional states of the elderly. This work presents a non-linear model for discriminating different arousal levels through quadratic entropy and a decision tree-based algorithm. Two hundred and seventy eight EEG recordings lasting one minute each were used to train the proposed model. The recordings belong to the Dataset for Emotion Analysis using Physiological signals (DEAP). In agreement with the complexity and variability observed in other works, our results report a low quadratic entropy when subjects face high arousal stimuli. Finally, the model achieves a global performance around 70 % when discriminating between calm and excitement events.

**Keywords** Entropy · EEG · Elderly · Monitoring · Modelling

## 1 Introduction

Nowadays, developed countries are experiencing a dramatic increase in population ageing produced mainly by a raise in life expectancy [1]. According to United Nations, the percentage of world population over 65 is about 9 % in 2016 and it is expected to increase up to 16 % in 2050 [2]. Among others, advances in medical services and improvements of healthcare systems are some reasons for this increasing

---

A. Martínez-Rodrigo (✉) · R. Alcaraz · B. García-Martínez · R. Zangróniz  
Instituto de Tecnologías Audiovisuales, Universidad de Castilla-La Mancha,  
13071 Cuenca, Spain  
e-mail: arturo.martinez@uclm.es

A. Fernández-Caballero  
Instituto de Investigación en Informática de Albacete,  
Universidad de Castilla-La Mancha, 02071 Albacete, Spain

trend. Unfortunately, the economic and social effects of an ageing population are considerable; and this phenomenon is threatening the balance on pensions and social systems [3]. Thus, healthcare is the largest area of expenditure in developed countries, increasing dramatically the cost as population ages. Bearing this in mind, it is not surprising that the World Health Organization considers cost reduction in healthcare systems as one of the most promising challenges in the upcoming years [4].

It is well known that elderly people generally prefer to live in their own homes over other options [5, 6]. This fact causes an additional tear in the healthcare system balances since elderly have to be attended at their homes. In order to palliate this negative effect, an important number of works have been performed, focusing on research and development of telemedicine equipments, fitness accessories and entertainment systems. The main objective of these works is to improve the well-being of elderly who decide to stay at their homes, but always from a physical point of view. To this respect, few efforts have been made at trying to regulate the elderly's mental and emotional state, even though it is known that mood greatly affects the elderly welfare [7]. Indeed, it has been reported that elderly people become more reclusive as loneliness sets in, increasing the probability of suffering mental illnesses like depression, anxiety or stress [8]. Similarly, it is also reported on the high incidence of elderly people suffering from disorientation or delirium while staying at home, which reveals an underlying mental disorder [9]. Taking this into account, emotion detection is the first step to develop systems based on emotional intelligence. Such systems could help us to detect some basic emotions like stress or calm, which has a vital role for improving well-being of elderly at home.

Within this context, it seems that technology should play a key role to face this challenge by innovating in healthcare. The use of assistive technologies, together with ubiquitous and pervasive computing, can help to palliate the ageing phenomenon. Novel advances in electronic devices, sensor miniaturization, power saving and development of power efficient communication protocols, among others, have led to the development of wireless body area networks (WBAN). WBAN consist of sensor networks that are worn on clothes or implanted on human body, allowing a continuous monitoring of physiological signals. Among physiological signals that are integrated into WBAN, there is electroencephalograph (EEG) registers. Recent studies have started to measure directly the electrical activity produced on the scalp by means of EEG, because cognitive processes such as emotions are primarily generated in our brain [10].

This paper studies the differences between two opposite levels of arousal, i.e. the excitement and calmness that a stimulus produces, throughout a standard EEG 10-20 electrodes system by using a non-linear methodology based on the quadratic entropy. The main objective is to create and evaluate the performance of a mathematical model capable of discriminating between these two emotional states.

## 2 Measuring Emotional States

From a psycho-physiological point of view, emotion is a mental process characterized by a strong activity and high degree of hedonistic content, resulting in physical and psychological changes that influence behaviour [11]. Different people feel distinct emotions when facing the same stimuli, depending on their mood, personality, disposition and motivation. However, regardless of the subjects, neuro-physiological researchers agree that feeling an emotion leads to changes in the physiological activity [12]. The physiological alterations caused by emotions produce variations in the arousal of the nervous system with different levels of depth, which seems to be related to different emotional states [13]. Traditionally, research on emotions has been carried out by studying physical aspects such as speech or facial expression. However, these features introduce uncertainty due to differences between facial expression and tonal speech in different cultures [14]. In contrast, physiological signals are more robust against these discordances, because they are connected directly with the autonomic central nervous system. Thus, heart rate variability, electro-dermal activity, electromyogram and skin temperature are some of the variables widely used in the last years [15]. However, recent studies have started to measure directly the electrical activity produced on the scalp by means of EEG registers. In this regard, it seems more appropriate to directly measure in the main source of the emotional processes, rather than through secondary variables produced indirectly by processes happening in the brain (sweat, muscle contractions, breathing increasing, etc.).

Nevertheless, measuring EEG data involves an intensification in signal processing methods, since non-linearity nature and complex dynamics of EEG signals require to use non-linear methods [16]. Indeed, the difficulty to detect changes in physiological EEG features through traditional lineal methods has been reported [17]. Otherwise, EEG recording approach presents some advantages over the rest of bio-signals due to the standardization of the electrode placement on the scalp (10-20 system), guaranteeing the reproducibility of the experiments and the comparison among different studies. However, measuring emotions is not a trivial task since a number of cognitive processes are evolved at the same time. In fact, a complex emotion consists of the combination of different feelings. Indeed, the lack of emotional intelligence in human-computer interaction systems has been widely reported, just because machines are inefficient in recognizing human emotional states. Therefore, research in predictive models able to understand basic emotions is only the first step in the future creation of more complex systems that will adapt machines to humans needs.

### 3 Methodology

#### 3.1 Database Description

The Dataset for Emotion Analysis using Physiological Signals (DEAP) [18] has been chosen, as it contains data collected from people with different arousal levels. This publicly available multi-modal database covers data from 32 participants, where each one watched 40 one-minute long excerpts from music videos. Then, participants rated each video in terms of arousal and valence by using a self-assessment manikin (SAM). 32-channels EEG signals were recorded and post-processed for each participant. The signals were high-pass filtered with a 2 Hz cut-off frequency and eye artefacts were removed by means of a blind source separation technique. Additional information about the details of this database are publicly available [18].

The high number of samples forming this database and the use of audiovisual stimuli makes this set suitable for our proper experimentation. In this regard, it has been reported that audiovisual segments are comparable to naturally occurring emotions, thus being more appropriate to elicit emotions [19]. From the entire database, an EEG recordings subset was formed with those segments where subjects reported an arousal level lower than 3 (group A or calmness) and higher than 7 (group B or excitement). Finally, the group under study consisted of 278 samples, where 147 segments belong to group A and 131 to group B, respectively.

#### 3.2 Definition of Quadratic Sample Entropy

Sample Entropy (*SampEn*) examines a time series for similar epochs and assigns a non-negative number to the sequence, with larger values corresponding to more irregularity in the data [20]. Two input parameters, a run length  $m$  and a tolerance window  $r$ , must be specified for *SampEn* to be computed.  $\text{SampEn}(m, r, N)$ , being  $N$  the length of the time series, is the negative logarithm of the conditional probability that two sequences similar for  $m$  points remain similar at the next point, where self-matches are not included in calculating the probability. Thus, a lower value of *SampEn* also indicates more self-similarity in the time series.

Formally, given  $N$  data points from a time series  $\{x(n)\} = x(1), x(2), \dots, x(N)$ , *SampEn* can be defined as follows [20]:

1. Form vector sequences of size  $m$ ,  $\mathbf{X}_m(1), \dots, \mathbf{X}_m(N - m + 1)$ , defined by  $\mathbf{X}_m(i) = \{x(i), x(i + 1), \dots, x(i + m - 1)\}$ , for  $1 \leq i \leq N - m$ . These vectors represent  $m$  consecutive  $x$  values, starting with the  $i$ th point.
2. Define the distance between vectors  $\mathbf{X}_m(i)$  and  $\mathbf{X}_m(j)$ ,  $d[\mathbf{X}_m(i), \mathbf{X}_m(j)]$ , as the absolute maximum difference between their scalar components,

$$d[\mathbf{X}_m(i), \mathbf{X}_m(j)] = \max_{k=0, \dots, m-1} \left( |x(i+k) - x(j+k)| \right). \quad (1)$$

3. For a given  $\mathbf{X}_m(i)$ , count the number of  $j$  ( $1 \leq j \leq N - m$ ,  $j \neq i$ ), denoted as  $B_i$ , such that the distance between  $\mathbf{X}_m(i)$  and  $\mathbf{X}_m(j)$  is less than or equal to  $r$ . Then, for  $1 \leq i \leq N - m$ ,

$$B_i^m(r) = \frac{1}{N - m - 1} B_i. \quad (2)$$

4. Define  $B^m(r)$  as

$$B^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_i^m(r). \quad (3)$$

5. Increase the dimension to  $m + 1$  and calculate  $A_i$  as the number of  $\mathbf{X}_{m+1}(i)$  within  $r$  of  $\mathbf{X}_{m+1}(j)$ , where  $j$  ranges from 1 to  $N - m$  ( $j \neq i$ ). Then,  $A_i^m(r)$  is defined as

$$A_i^m(r) = \frac{1}{N - m - 1} A_i. \quad (4)$$

6. Set  $A^m(r)$  as:

$$A^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_i^m(r). \quad (5)$$

Thus,  $B^m(r)$  is the probability that two sequences will match for  $m$  points, whereas  $A^m(r)$  is the probability that two sequences will match for  $m + 1$  points. Finally, *SampEn* can be defined as

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \left\{ -\ln \left[ \frac{A^m(r)}{B^m(r)} \right] \right\}, \quad (6)$$

which is estimated by the statistic

$$\text{SampEn}(m, r, N) = -\ln \left[ \frac{A^m(r)}{B^m(r)} \right]. \quad (7)$$

Although  $m$  and  $r$  are critical in determining the outcome of *SampEn*, no guidelines exist for optimising their values. In principle, the accuracy and confidence of the entropy estimate improve as the number of length  $m$  matches increases. The number of matches can be increased by choosing small  $m$  (short templates) and large  $r$  (wide tolerance). However, penalties appear when too relaxed criteria are used [21]. For smaller  $r$  values, poor conditional probability estimates are achieved, while for larger  $r$  values, too much detailed system information is lost and *SampEn* tends to 0 for all processes. To avoid a significant noise contribution on *SampEn* computation, one must choose  $r$  larger than most of the noise [21]. A slight modification of *SampEn* to do it more insensitive to the  $r$  selection is the named *Quadratic*

*SampEn* (QSE) [22]. This new measure allows to vary  $r$  as needed to achieve confident estimates of the conditional probability and is defined as

$$QSE(m, r, N) = SampEn(m, n, r) + \ln(2r). \quad (8)$$

Anyway, it is worth noting that the most widely established values for *SampEn* computation are  $m = 2$  and  $r = 0.25$  times the standard deviation of the original data [23] and, therefore, they were used in this work.

### 3.3 Statistical Analysis

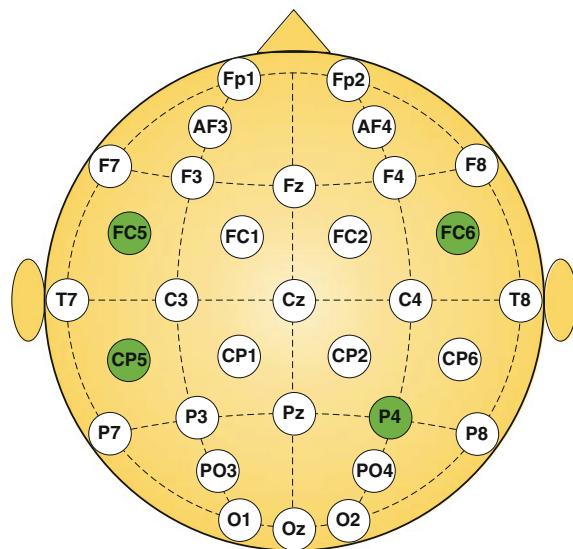
Shapiro-Wilks and Levene test proved that distributions were normal and homoscedastic. Consequently, results are expressed as mean  $\pm$  standard deviation for all the samples belonging the same group. Statistical differences for the two groups and for each channel were assessed by a one-way ANOVA test. In this regard, a value of statistical significance  $\rho < 0.05$  was considered as significant. To evaluate the discriminatory power for each recorded EEG channel, receiver operating characteristic curves (ROC) were used. Brief, discriminatory threshold between groups A and B were firstly computed, making use of all the segments described in the database. Then, optimal thresholds were chosen for each channel as those values of quadratic entropy reducing the classification error. Finally, total accuracy was calculated as the ratio between the properly classified samples and the total number of analysed samples. Moreover, to improve the classification performance between the two groups, a decision tree classifier was used to investigate the non-monotonic and non-linear relationships among the regularity of different channels. In this sense, the optimal combination of EEG channels was analysed. The stopping criterion for the tree growth is that each node contains only samples from one class or fewer than 20 % of all observations. The splitting criterion for each EEG channel, which is used to evaluate the goodness of the alternative splits for each attribute, was carried out by applying the Gini index [24].

## 4 Results

Table 1 shows the mean and standard deviation values of quadratic entropy computed from EEG channels for the two groups under study. Only those channels reporting statistical differences through a one-way ANOVA test are shown. As can be appreciated, 4 channels out of 32 reported statistical differences among the two groups. All of them are located in the frontal and parietal zones on the head, being symmetrical to each one (see Fig. 1). Thus, FC5 and FC6 correspond with left and right channels located in the fronto-central part of the brain, respectively. Similarly, CP5 and P4 correspond to left and right part of central-parietal and parietal zones on the brain,

**Table 1** Statistical significance between groups A and B

EEG channel	Group A	Group B	$\rho$
	Mean $\pm$ deviation	Mean $\pm$ deviation	
FC5	$3.1914 \pm 0.7667$	$2.9967 \pm 0.6836$	0.05
FC6	$3.2305 \pm 0.8798$	$2.9600 \pm 0.8881$	0.0253
CP5	$3.2442 \pm 0.7891$	$2.9484 \pm 0.7280$	0.0056
P4	$2.7824 \pm 0.7836$	$3.0721 \pm 0.6721$	0.0048

**Fig. 1** 10-20 EEG system representation

respectively. Information about the location and labelling of the different channels can be found in Fig. 1.

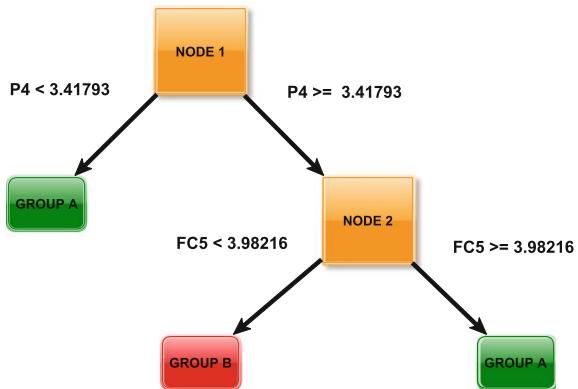
Results regarding average quadratic entropy show a decreasing complexity in FC5, FC6 and CP5 when the arousal level increases, while P4 presents an increasing trend when the arousal level arises. From a statistical point of view, the most remarkable trends are provided by parietal channels CP5 and P4. Regarding performance classification, Table 2 shows the results for each EEG channel by using a ROC approach. All the parameters present a comparable performance, achieving a global accuracy around 60 %. Nevertheless, P4 channel presents the highest predictive ability, reaching a global accuracy of 62.50 %, while FC5 presents the lowest classification performance, obtaining a global accuracy of 56.53 %.

Finally, in order to study the possible relationship among EEG channels, a decision tree algorithm is programmed. It is worth noting that all the channels are again included in the classification process. Then, the growing criteria and splitting rule are considered to compute the tree classification model. As can be appreciated, a two-level classification tree is obtained, where two parameters are chosen as the most

**Table 2** Discriminatory power of entropy for each EEG channel using a ROC curve and decision tree

<i>Receiver operating curve</i>			
Channel	Sensitivity (%)	Positive predictability (%)	Global accuracy (%)
<i>FC5</i>	58.72	54.55	56.53
<i>FC6</i>	49.54	75.76	62.02
<i>CP5</i>	61.47	57.58	59.43
<i>P4</i>	56.57	67.89	62.50
<i>Decision tree</i>			
<i>P4 and FC5</i>	63.29	82.00	<b>69.79</b>

**Fig. 2** Combination of P4 and FC5 channels by means of a decision tree classifier



representative to improve the overall performance classification (see Fig. 2). Thus, P4 is chosen first, and an entropy lower than 3.41793 is set as the threshold for those samples belonging to group A. Otherwise, the tree model is split again by using FC5 channel. In this case, an entropy lower than 3.98216 is established as the threshold for those samples belonging to group B, and entropy higher or equal than 3.98216 is set for those samples belonging to group A. Results using the classification tree can be observed in Table 2. As it is shown, the classification performance achieved by means of the combination of P4 and FC5 outperforms the results obtained with only P4. Thus, an improvement of more than 6 % and 14 % is observed for the sensitivity and positive predictability, respectively, improving the global accuracy more than 7 % regarding the analysis of P4 alone.

## 5 Discussion and Conclusions

In the last years, an intensification on the study of valence-arousal space from an EEG point of view is emerging as an alternative to the traditional methods for quantifying emotional states. The main reason is that brain signals represent actions directly from the source, while other physiological signals are activated by the autonomic nervous system in response to commands originated in the brain. In this work, a predictive non-lineal model to classify differences between low and high arousal level has been proposed. Although a low-high arousal level can contemplate positive and negative connotations depending on the valence level, our final application is focused exclusively on the excitement or calmness that a stimulus produces in the elderly.

Quadratic entropy is used in this research to determine the complexity or irregularity of EEG signals under different elicited arousal degrees. Given the statistical significance obtained for some channels, it seems that non-lineal analysis could be a good indicator of the brain working dynamics under different values of arousal. According to the results, excitement states experience lower entropy levels than calm states. These differences are mainly located in central and left parietal channels. The reduction of entropy values represents a decreasing in the brain system complexity and, therefore, an increasing of the regularity. On the contrary, when subjects are under calm events, brain complexity increases, arising the complexity of EEG signals. These findings are in line with other works where a reduction in fractal dimensions in participants under negative emotional stress has been reported [25]. Nevertheless, right parietal electrode P4 shows a different trend compared with the rest of significant channels. This finding could be related with the dissociation existing in the brain between arousal and valence. Indeed, it has been reported that changes in arousal and valence are linked with the brain activity at parietal location [26]. Given that our study is exclusively focused on changes in arousal, the quantification of how pleasant or unpleasant a stimulus is, could be influencing in the results obtained on P4.

Regarding classification performance, the proposed tree-based model yields a better diagnostic accuracy than when only one channel is considered. An improvement of more than 7 % is achieved by combining CP5 and P4, thus showing underlying relationships among different brain locations. On the other hand, only two channels out of thirty-two were considered in the model, thus improving an easier understanding of its outcome. Consequently, the implementation of this model in a real-time monitoring and diagnosis system is a future exploratory research line.

**Acknowledgments** This work was partially supported by Spanish Ministerio de Economía y Competitividad/FEDER under TIN2013-47074-C2-1-R and TIN2015-72931-EXP grants

## References

1. Lutz, W., Sanderson, W., Scherbov, S.: The coming acceleration of global population ageing. *Nature* **451**(7179), 716–719 (2008)
2. United Nations Department of Economic, World Population Ageing 2009, vol. 295. United Nations Publications (2010)
3. Carone, G., Costello, D., Diez Guardia, N., Mourre, G., Przywara, B., Salomäki, A.: The economic impact of ageing populations in the eu25 member states. Directorate General for Economic and Financial Affairs (236)
4. World Health Organization, et al.: Global Health and Ageing
5. Castillo, J.C., Castro-González, A., Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A., Salichs, M.A.: Software architecture for smart emotion recognition and regulation of the ageing adult. *Cognitive Computation*. In Press
6. Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A.: Improvement of the elderly quality of life and care through smart emotion regulation. In: *Ambient Assisted Living and Daily Activities*, pp. 348–355. Springer (2014)
7. Martínez-Rodrigo, A., Zangróniz, R., Pastor, J.M., Fernández-Caballero, A.: Arousal level classification in the ageing adult by measuring electrodermal skin conductivity. In: *Ambient Intelligence for Health*, pp. 213–223. Springer (2015)
8. Costa, Á., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. *Expert Syst. Appl.* **39**(15), 12192–12204 (2012)
9. García-Rodríguez, C., Martínez-Tomás, R., Cuadra-Troncoso, J.M., Rincón, M., Fernández-Caballero, A.: A simulation tool for monitoring elderly who suffer from disorientation in a smart home. *Expert Syst.* **32**(6), 676–687 (2015)
10. German, W.J.: The hypothalamus and central levels of autonomic function. *Yale J. Biol. Med.* **12**(5), 602–603 (1940)
11. Schacter, D.L.: *Psychology*, 2nd edn. (2011)
12. Nasoz, F., Lisetti, C.L., Alvarez, K., Finkelstein, N.: Emotion recognition from physiological signals for user modeling of affect. In: *Proceedings of the 3rd Workshop on Affective and Attitude User Modelling*, pp. 1–8 (2003)
13. Russell, J.A.: A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**(6), 1161 (1980)
14. Kim, J.: Bimodal emotion recognition using speech and physiological changes. In: *Robust Speech Recognition and Understanding*, INTECH Open, pp. 265–280 (2007)
15. Valenza, G., Lanata, A., Scilingo, E.P.: The role of nonlinear dynamics in affective valence and arousal recognition. *IEEE Trans. Affect. Comput.* **3**(2), 237–249 (2012)
16. Hatamikia, S., Nasrabadi, A.: Recognition of emotional states induced by music videos based on nonlinear feature extraction and som classification. In: *21th Iranian Conference on Biomedical Engineering*, pp. 333–337. IEEE (2014)
17. Akar, S.A., Kara, S., Agambayev, S., Bilgic, V.: Nonlinear analysis of eeg in major depression with fractal dimensions. In: *37th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society*, pp. 7410–7413. IEEE (2015)
18. Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: Deep: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* **3**(1), 18–31 (2012)
19. Philippot, P.: Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cogn. Emot.* **7**(2), 171–193 (1993)
20. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circulatory Physiol.* **278**(6), 2039–2049 (2000)
21. Pincus, S.M.: Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci.* **88**(6), 2297–2301 (1991)
22. Lake, D.E., Moorman, J.R.: Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Physiol. Heart Circulatory Physiol.* **300**(1), 319–325 (2011)

23. Alcaraz, R., Abásolo, D., Hornero, R., Rieta, J.J.: Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. *Comput. Methods Prog. Biomed.* **99**(1), 124–132 (2010)
24. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and Regression Trees. CRC Press (1984)
25. Hosseini, S.A., Naghibi-Sistani, M.B.: Classification of emotional stress using brain activity. In: Applied Biomedical Engineering, INTECH Open, pp. 313–336 (2011)
26. Dolcos, F., Cabeza, R.: Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. *Cogn. Affect. Behav. Neurosci.* **2**(3), 252–263 (2002)

# Emotional Induction Through Films: A Model for the Regulation of Emotions

**Luz Fernández-Aguilar, José Miguel Latorre, Laura Ros,  
Juan Pedro Serrano, Jorge Ricarte, Arturo Martínez-Rodrigo,  
Roberto Zangróniz, José Manuel Pastor, María T. López and  
Antonio Fernández-Caballero**

**Abstract** This paper introduces a software program to recognise discrete emotions on an ageing adult from his/her physiological and psychological responses. This research considers the capacity from an audiovisual method to evoke different emotions and uses it to interpret and modulate basic emotion states. Different body sensors, in the case of physiological response, and a set of questionnaires, in the case of psychological responses, are selected to measure the power in causing fear, anger, disgust, sadness, amusement, affection and the neutral state, through a set of films used as an emotional induction method. The initial results suggest that it is possible to extract discrete values about positive and negative emotional states with films and to use these responses as keys to get emotion regulation.

**Keywords** Emotion induction · Films · Emotional regulation

## 1 Introduction

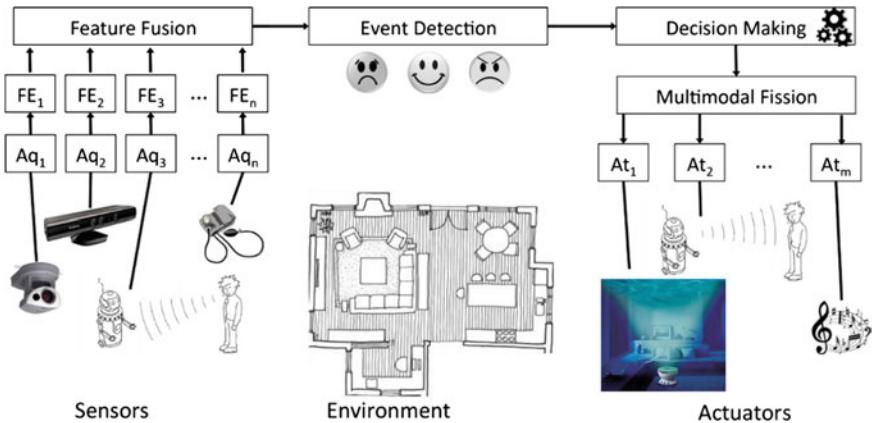
In the history of health research, numerous papers have focused on physical aspects, especially in older people. Although the concept of quality of life includes “physical health of the subjects, their psychological state, level of independence, social relationships, and their relationship with the environment” [1], only a few works have attempted to regulate their emotional state. The recognition of emotions is very

---

L. Fernández-Aguilar (✉) · J.M. Latorre · L. Ros · J.P. Serrano · J. Ricarte  
Universidad de Castilla-La Mancha, Unidad de Psicología Cognitiva Aplicada (CICYPA),  
02071 Albacete, Spain  
e-mail: MariaLuz.Fernandez@uclm.es

A. Martínez-Rodrigo · R. Zangróniz · J.M. Pastor  
Instituto de Tecnologías Audiovisuales, Universidad de Castilla-La Mancha,  
13071 Cuenca, Spain

M.T. López · A. Fernández-Caballero (✉)  
Instituto de Investigación en Informática de Albacete, Universidad de Castilla-La Mancha,  
02071 Albacete, Spain  
e-mail: Antonio.Fdez@uclm.es



**Fig. 1** Outline of the project

important for the relationship between people, in this case ageing adults, and their psychological health.

On the one hand, thanks to technological and medical advances, there has been an increase in life expectancy and an increasing ageing population [2]. On the other hand, most people prefer to stay at home as long as their quality of life does permit it. From both premises the project “Improvement of the Elderly Quality of Life and Care through Smart Emotion Regulation” [3–5] is born (see Fig. 1). It addresses a challenge related to “Economy and Digital Society”, linked to the Spanish Strategy for Science, Technology and Innovation. The main goal of this project is to look for tools capable of improving the quality of life and care of the ageing adult who continues living at home by choice. This has to be achieved through the creation/use of Ambient Intelligence and smart emotion regulation devices. Indeed, we believe that the ability to monitor changes in the emotional state of a person in his/her own context allows implementing regulatory strategies for reducing negative affect. In some specific goals, we want to:

1. Develop a system capable of detecting emotions through facial expression and physiological response, and adapt the system for use with elderly.
2. Develop an emotion regulation system by using stimuli like colour, lighting, sound landscape, music or autobiographical memories, among others, and adapt the system for use with ageing adults.

To get all of this, cameras and body sensors are used for monitoring the older people’s facial and gestural expression, activity and behaviour, as well as relevant physiological data [6, 7]. This way, the older people’s emotions are inferred and recognised. Music, colour and light are the means of stimulation to regulate their emotions towards a positive and pleasant mood [8, 9]. The present research focuses on the use of films as method to detect basic emotions.

## 2 Films and Emotion Recognition

For research it is essential to have effective screening models of emotions that allow their application in the field of emotion regulation. On the one hand, mood induction procedures (MIPs) are making a great contribution to our understanding of the relationship between what we feel, what we think and what we do, to our understanding the links between emotion, cognition and behaviour in normal and clinical populations [10]. On the other hand, Psychology has studied a large number of these methods to provoke emotions differentially, or what is the same, there are different ways to manipulate emotional states in experimental research designs as pictures, music, imagination, Velten elicitation or films [11]. For the purpose of our work, we have selected a set of films that should act as stimuli capable of eliciting the basic emotions (anger, fear, amusement, affection, sadness and disgust) besides of the neutral state [12].

Induction by films consists in watching a brief set of fragments which usually come from commercial films. These fragments have an intense content of positive or negative emotions. At present it is one of the most widely used techniques due to films have a greater ecological validity because they promote a dynamic context with auditory and visual stimuli similar to those that can be seen in real life. In addition, they are a standardized method that does not require individual adaptations as other as hypnosis or imagination procedures which involving subjective techniques [13]. It has also been shown that the use of film scenes for induction does not require large attentional effort and that when the instructions to evaluate them are not too direct, the potential problem that subjects respond differently because it is an experimental context decreases (demand effect).

Films are a method and there exists an evidence that they elicit the activations across many of the response systems associated with emotions. Indeed, films can elicit discrete emotions unlike other methods as music or smells that can only differentiate between pleasant and unpleasant [7, 13–15]. They are also preferred due to their degree of standardization and the possibility to use them with relatively low levels of demand [13, 14, 16]. Moreover, films allow an accurate selection of the stimuli according to their position in the affective space defined by the dimensions of valence, arousal and dominance. In the current study, these variables are evaluated from their physiological and psychological responses.

To measure the subjective responses, we use different psychological tests:

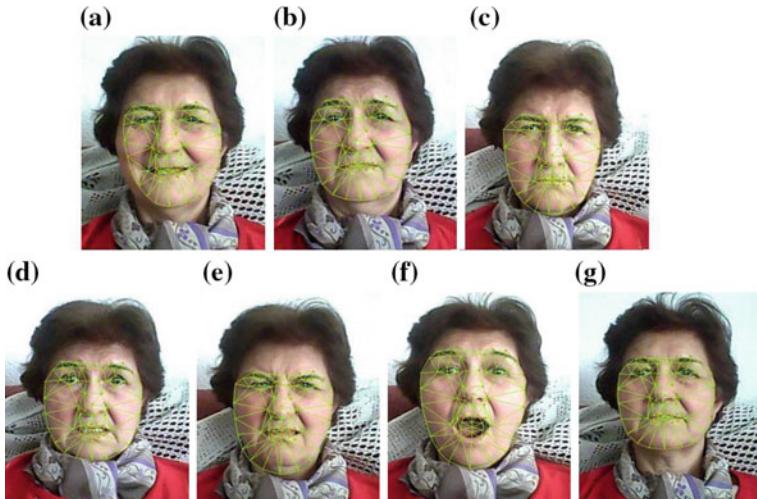
- *Self Assessment Manikin (SAM)* [17]: This test evaluates the dimensions of valence, arousal and dominance through five figures or manikins by every dimension (Likert scale 1–9).
- *Discrete Emotions Questionnaire (CED)* (Spanish version) [14]: There are 18 items with emotional labels (Likert scale 1–7): fun, rage/anger, anxiety, confusion, satisfaction, disgust, fear, guilt, happiness, interest, joy, love/affection, pride, sadness, shame, surprise, unhappiness and embarrassment.

- Control questions (YES/NO): The participant is asked to answer some question sometime during some fragment, and whether he/she has ever seen the film previously.

To measure the objective responses, we evaluate different physiological responses with cameras and sensors. In particular, a wristband is used that includes:

- *Electrodermal activity* (EDA): The spontaneous skin conductance (SSC) is the result of an increasing activity in the sympathetic nervous system [18], and the basal skin conductance (BSC) is related both with the sympathetic nervous system and the dermal characteristics of skin [19]. Different markers that evaluate the intensity and duration of the events are computed [20].
- *Heart activity*: It is caused by the autonomous nervous system. Heart rate variability (HRV), showing the alterations of heart rhythm, is usually computed to evaluate the arousal level of an individual [21].
- *Temperature*: If human body is under stress, his/her temperature drops due to the contraction of blood vessels [22].

Also, *facial emotion detection* is provided. We use a real-time facial expression recognition system based on geometric features. This system works by detecting facial points associated with emotions and making a classification for each emotional category [23, 24] (see Fig. 2).



**Fig. 2** Example of webcam capture where the detected emotion is **a** Joy. **b** Sadness. **c** Anger. **d** Fear. **e** Disgust. **f** Surprise. **g** Neutral

### 3 Participants and Procedure for Emotion Elicitation by Films

The participants consisted of fifteen persons (seven females and eight males). They participated voluntarily and received no financial compensation. All participants were between the ages 18 and 83 years. Young subjects were people under 35 years old ( $M = 20.16$ ), participants between 35 and 59 years old were considered middle-age ( $M = 46.83$ ) and subjects over than 60 years of age were considered elderly ( $M = 71.66$ ), as you can appreciate in Table 1. All participants were recruited from Universidad de Castilla-La Mancha (students and workers) and from University of Experience (academic courses for elderly). The three age groups had similar years of education. No participants suffered from severe chronic illness, neurological and/or mood disorders according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria. Moreover, none of the participants were taking any medication that could affect the process or task results.

The experiment is performed in a small room equipped with a comfortable armchair and a 27-inches screen monitor TV. Upon arrival of the participants, they are welcomed by offering an overview of the experiment and the sign a written consent. Before starting the experimental session, Beck Depression Inventory (BDI) [25], and Positive and Negative Affect Schedule (PANAS) [26], are administered to know the participant's current emotional state. Moreover, if the participant is older than 60, Mini-Mental State Examination (MMSE) [27] is administered to rule out any cognitive impairment.

The experimental task has an average duration of 50 min depending on the test participant who may answer the complete questionnaire quicker or slower. The experiment has been designed with software E-Prime 2.0, which includes, for each event, the instructions for the experiment, the bank of audiovisual stimuli composed of 55

**Table 1** Means (and standard deviations) of valence, arousal and dominance ratings from the Self Assessment Manikin (SAM) for the neutral, negative and positive films

SAM (n = 15)	Valence M (SD)	Significance p	Arousal M (SD)	Significance p	Dominance M (SD)	Significance p
(Rank 1–9)						
Neutral	6.20 (1.61)		3.90 (2.30)		6.86 (2.00)	
Anger	1.93 (1.87)	0.000	6.67 (2.16)	0.000	4.33 (2.25)	0.006
Sadness	3.93 (2.86)	0.011	5.93 (1.98)	0.004	5.60 (2.38)	0.186
Fear	2.67 (1.71)	0.000	6.53 (2.20)	0.000	5.33 (1.29)	0.032
Disgust	2.27 (2.12)	0.000	5.87 (2.20)	0.000	5.67 (1.91)	0.124
Affection	6.73 (2.28)	0.310	5.07 (2.18)	0.012	5.33 (2.28)	0.093
Amusement	7.73 (1.83)	0.020	5.07 (2.28)	0.029	6.20 (1.78)	0.232

*Note* T-TEST (Significance): statistically significant differences between the neutral films and other emotional categories. Critical value of  $p < 0.050$

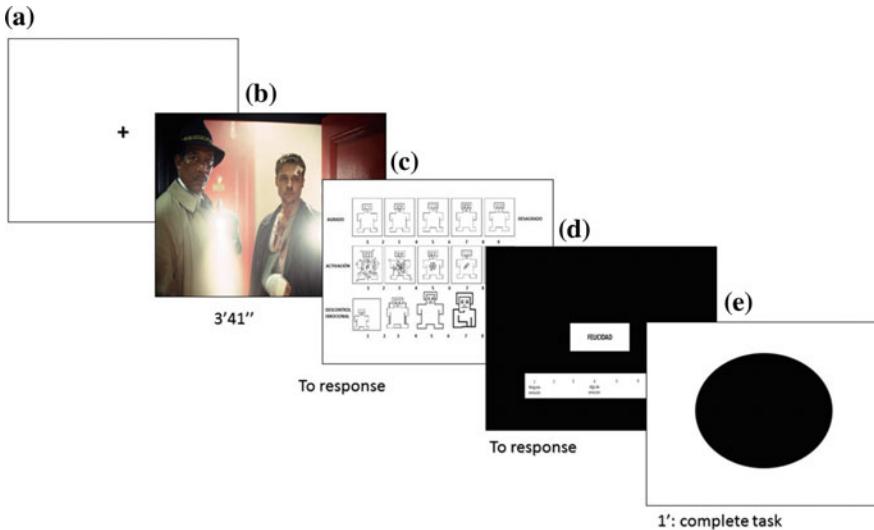
NEUTRAL	AMUSEMENT	AFFECTION	DISGUST	ANGER	FEAR	SADNESS
Blue (1)	When Harry met Sally	La vita è bella (2)	Pink Flamingos	Schindler's list (2)	Scream (2)	City of angels
Blue (2)	There's Something About Mary (1)	La vita è bella (3)	Trainspotting (2)	Schindler's list (3)	The Blair witch project	The Champ
L'amant	There's Something About Mary (2)	La vita è bella (4)	Hellraiser	American History X	The exorcist	Schindler's list (1)
	The dinner game	Forrest Gump	Seven (3)	In the name of the father	The silence of the lambs (2)	Dangerous minds
	Dumb and Dumber	When a Man Loves a Woman	The dentist	Leaving Las Vegas	Seven (2)	A perfect world
	Benny Joone	Dead Poets Society (2)	27 days	Cry freedom	The shining	La vita è bella (1)
	Les visiteurs	Leon	Saving private Ryan	Sleepers	Misery	Philadelphia
	A fish called Wanda	E.T	The silence of the lambs(1)	Seven (1)	Copycat	La vie revée des anges
		Ghost		The piano		Dead man walking
						Dead Poets Society (1)

**Fig. 3** Classification of films according to discrete emotions

films (ranging between 24 s, and 6 min and 4 s) [13, 14, 27] relevant questionnaires and a distracting task. The films used to the experiment are listed in Fig. 3.

At the beginning of the experiment, the electrodes are positioned and the camera is activated while the experiment procedure is explained. First, the computer presents basic instructions for starting the experiment. Next, two neutral film clips are presented so that the participant can practice with the task. As he/she is relaxed, we get a baseline of the different physiological parameters. After seeing a scene, the participant must complete the SAM and the CED questionnaires for each scene [12–14] and perform a distracting task for one minute. The distracting task is the emergence of a geometric figure on the screen for 5000 ms. When the figure is a circle the participant must press 1, and when a different figure is shown on the screen the participant must press 2. This task is offered to prevent the cumulative effect of one emotion to another. In total, every subject watches a set of 9 films including 7 emotional fragments that appear in a counterbalanced way.

At the end of the session, another neutral clip is presented to recover a relaxing state. Physiological parameters and facial expressions can be evaluated at the same time that the participants watch each film clip, and the subjective emotional response is evaluated immediately after each film clip. A example of an event sequence is shown in Fig. 4.



**Fig. 4** Sequence of one event

## 4 Initial Results and Conclusions

As described before, fifteen volunteers were recruited at Albacete School of Medicine, Universidad de Castilla-La Mancha, Spain, to take part in the pilot test. The initial results suggest that films are a good method to elicit different emotion states. Also, the wristband and the camera are good tools to measure key parameters in basic emotions. Although the number of participants is small to draw clear conclusions, we believe that there is some evidence about general tendencies related to the results obtained from physiological and psychological responses.

This article has described the first steps in the use of films to induce emotions. The objective of this study is to find solutions for regulating affect and thus improving the quality of life and care of ageing adults living at home. The initial results are complying with the goals of our running project related to the improvement of the elderly quality of life. For further progress, it is expected to significantly increase both the sample size and the inclusion of older people in the near future.

**Acknowledgments** This work was partially supported by Spanish Ministerio de Economía y Competitividad/FEDER under TIN2013-47074-C2-1-R grant.

## References

1. World Health Organization. Ageing and Life Course (2011)
2. Castillo, J.C., Fernández-Caballero, A., Castro-González, Á., Salichs, M.A., López, M.T.: A framework for recognizing and regulating emotions in the elderly. In: Ambient Assisted Living and Daily Activities, pp. 320–327 (2014)
3. Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A.: Improvement of the elderly quality of life and care through smart emotion regulation. In: Ambient Assisted Living and Daily Activities, pp. 348–355 (2014)
4. Costa, A., Castillo, J.C., Novais, P., Fernández-Caballero, A., Simoes, R.: Sensor-driven agenda for intelligent home care of the elderly. *Expert Syst. Appl.* **39**(15), 12192–12204 (2012)
5. Castillo, J.C., Castro-González, Á., Fernández-Caballero, A., Latorre, J.M., Pastor, J.M., Fernández-Sotos, A., Salichs, M.A.: Software architecture for smart emotion recognition and regulation of the ageing adult. *Cogn. Comput.* (2016). doi:[10.1007/s12559-016-9383-y](https://doi.org/10.1007/s12559-016-9383-y)
6. Fernández-Caballero, A., Castillo, J.C., Rodríguez-Sánchez, J.M.: Human activity monitoring by local and global finite state machines. *Expert Syst. Appl.* **39**(8), 6982–6993 (2012)
7. Lench, H.C., Flores, S.A., Bench, S.W.: Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychol. Bull.* **137**(5), 834–855 (2011)
8. Fernández-Sotos, A., Fernández-Caballero, A., Latorre, J.M.: Elicitation of emotions through music: the influence of note value. In: Artificial Computation in Biology and Medicine, pp. 488–497 (2015)
9. Ortiz-García-Cervigón, V., Sokolova, M.V., García-Muoz, R., Fernández-Caballero, A.: LED strips for color- and illumination-based emotion regulation at home. In: Ambient Assisted Living. Development and Testing of ICT-Based Solutions in Real Life Situations, pp. 277–287 (2015)
10. Salas, C.E., Radovic, D., Turnbull, O.H.: Inside-out: comparing internally generated and externally generated basic emotions. *Emotion* **12**(3), 568–578 (2012)
11. Ekman, P.: Expression and the nature of emotion. In: Scherer, K.R., Ekman, P. (eds.) Approaches to Emotions. Erlbaum, Hillsdale (1984)
12. Rottenberg, J., Ray, R.D., Gross, J.J.: Emotion elicitation using film. In: Coan, J.A., Allen, J.B. (eds.) Handbook of Emotion Elicitation and Assessment, pp. 9–28. Oxford University Press, New York (2007)
13. Fernández, C., Pascual, J.C., Soler, J., Elices, M., Portella, M.J., Fernández-Abascal, E.: Physiological responses induced by emotion-eliciting films. *Appl. Psychophysiol., Biofeedback* **37**(2), 73–79 (2012)
14. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**(1), 49–59 (1994)
15. Philippot, P., Schaefer, A., Herbette, G.: Consequences of specific processing of emotional information: impact of general versus specific autobiographical memory priming on emotion elicitation. *Emotion* **3**, 270–283 (2003)
16. Schaefer, A., Fletcher, K., Pottage, C., Alexander, K., Brown, C.: The effects of emotional intensity on ERP correlates of recognition memory. *NeuroReport* **20**(3), 319–324 (2009)
17. Lidberg, L., Wallin, G.: Sympathetic skin nerve discharges in relation to amplitude of skin resistance responses. *Psychophysiology* **18**(3), 268–270 (1981)
18. Venables, P.H., Christie, M.J.: Electrodermal activity. In: Techniques in Psychophysiology, pp. 3–67 (2012)
19. Leijdekkers, P., Gay, V., Frederick, W.: CaptureMyEmotion: a mobile app to improve emotion learning for autistic children using sensors. In: 26th IEEE International Symposium on Computer-Based Medical Systems, pp. 381–384 (2013)
20. Martínez-Rodrigo, A., Zangróniz, R., Pastor, J.M., Fernández-Caballero, A.: Arousal level classification in the ageing adult by measuring electrodermal skin conductivity. In: Ambient Intelligence for Health, pp. 213–223 (2015)

21. Martínez-Rodrigo, A., Zangróniz, R., Pastor, J.M., Latorre, J.M., Fernández-Caballero, A.: Emotion detection in ageing adults from physiological sensors. In: Ambient Intelligence-Software and Applications, pp. 253–261 (2015)
22. Zhou, Q., Wang, X.: Real-time facial expression recognition system based-on geometric features. *Lect. Notes Electr. Eng.* **212**, 449–456 (2013)
23. Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J.: An inventory for measuring depression. *Arch. Gen. Psychiatry* **4**, 561–571 (1961)
24. Lozano-Monasor, E., López, M.T., Fernández-Caballero, A., Vigo-Bustos, F.: Facial expression recognition from webcam based on active shape models and support vector machines. In: Ambient Assisted Living and Daily Activities, pp. 147–154 (2014)
25. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* **54**, 1063–1070 (1988)
26. Folstein, M., Folstein, S.E., McHugh, P.R.: “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**(3), 189–198 (1975)
27. Schaefer, A., Nils, F., Sanchez, X., Philippot, P.: Assessing the effectiveness of a large database of emotion-eliciting films: a new tool for emotion researchers. *Cogn. Emot.* **24**(7), 1153–1172 (2010)

# **Application of the Lognormal Model to the Vocal Tract Movement to Detect Neurological Diseases in Voice**

**Cristina Carmona-Duarte, Réjean Plamondon, Pedro Gómez-Vilda,  
Miguel A. Ferrer, Jesús B. Alonso and Ana Rita M. Londral**

**Abstract** In this paper a novel method to evaluate the quality of the voice signal is presented. Our novel hypothesis is that the first and second formants allow the estimation of the jaw-tongue dynamics. Once the velocity is computed, it is approximated by the Sigma-Lognormal model whose parameters enable to distinguish between normal and pathological voices. Three types of pathologies are used to test the method: Laryngeal Diseases, Parkinson and Amyotrophic Lateral Sclerosis. Preliminary results show that the novel features proposed are able to distinguish between parameters of normal and pathological voice. Moreover, it is also possible to discriminate between the three types of pathologies studied in this work.

---

C. Carmona-Duarte (✉) · M.A. Ferrer · J.B. Alonso

Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones,  
Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain  
e-mail: ccarmona@idetic.eu; criscarmonad@gmail.com

M.A. Ferrer

e-mail: mferrer@idetic.eu

J.B. Alonso

e-mail: jalonso@idetic.eu

R. Plamondon

Département de Génie Électrique, Laboratoire Scribens, École Polytechnique de Montréal,  
Montréal, Canada  
e-mail: rejean.plamondon@polymtl.ca

P. Gómez-Vilda

Facultad de Informática, Universidad Politécnica de Madrid,  
Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Madrid, Spain  
e-mail: pedro@fi.upm.es

A.R.M. Londral

Translational Clinical Physiology Unit, Instituto de Medicina Molecular,  
Universidade de Lisboa, Lisbon, Portugal

**Keywords** Sigma-lognormal • Formant • Kinematic • Articulation • Parkinson • ALS

## 1 Introduction

The Kinematic Theory of rapid human movements [1] has successfully been applied to handwriting modeling. According to this theory, the way in which neuromuscular systems are involved in the production of muscular movements is modelled using lognormal velocity profiles, here after referred to as the Sigma-lognormal model [1]. This theory has been successfully used in handwriting, studying the neuromuscular system for rapid movements [2], the handwriting variations in time [3, 4], the specification of a diagnostic system for neuromuscular disorder [5], the assessment of brain stroke risk factors [6], etc.

Just as the trunk and arm muscles act to generate handwriting, in the case of speech articulation, the tongue and jaw muscles move to generate different vowel sounds.

In a preliminary study [7] we have investigated the possibility of applying the Kinematic Theory of rapid human movements to analyse the acoustic-phonetic articulation. This study validated the hypothesis that vocal tract dynamics, which includes jaw and tongue could be modelled by Sigma-Lognomals as calculated from the velocity profile from speech signal.

Indeed, in speech production, the resonating cavities modifiable by the articulatory organs allow the energy of the speech signal to be focused at certain frequencies (formants), due to oropharyngeal tract resonators. In speech analysis, it is known that a specific vowel phonation is related to the velopharyngeal switch, the mandibular system, the tongue neuromotor system and the laryngeal system. When the vowels are analysed, it is possible to obtain from the two first formants (F1 and F2) [8] a representation space as described in [9]. Specifically, the Kinematic Theory models the movement of tongue dynamics by the velocity computed by a transformation of the first and second formant frequencies.

There are many studies dealing with pathology detection by mean of the speech processing such as: Laryngeal Diseases, Parkinson and Amyotrophic Lateral Sclerosis.

In the case of abnormal vocal fold structure or noise in the speech recorded, formant extraction could fail producing artefacts; in this case the lognormal reconstruction would show low Signal to Noise Ratio (SNR) due to an incorrect change in formant extraction not resulting from a real movement. The pathological subjects in the Laryngeal Disease database presented abnormal vocal fold behavior [10, 11].

Parkinson's disease is usually produced by a dopamine deficiency [12] and it causes motor symptoms such as tremor, muscular rigidity, etc. [13]. Some studies showed that speech is one of the first biosignals affected [14]. There are recent

studies about evaluation of Parkinson speech using the variability of pitch, speech rate and pausing [15–18].

In turn, Amyotrophic Lateral Sclerosis is a neuromuscular disease characterized by a reduced velocity of speech movements affecting rhythm [19]. In this case, there are also previous studies about the evaluation of this disease using the speech signal [9, 20, 21].

In this work we applied the Kinematic Theory to a dataset of normal and laryngeal pathological speech. The objective is to parameterize healthy and pathological speech in order to find out if they can be automatically distinguished, and to establish different parameters within pathological speech patterns for subjects with or without neurological diseases.

In order to study the case of neurological pathologies, a database including three Parkinson and two Amyotrophic lateral sclerosis (ALS) subjects has been considered. On the other hand, to study laryngeal diseases, 20 laryngeal pathological and 51 healthy subjects were studied.

This paper is divided as follows: first a brief introduction explaining how the velocity is obtained from the formants is presented. Secondly, in Sect. 3, the kinematic model and the parameters used for speech evaluation are described. Thirdly, in Sect. 4 the method and subjects used in this work are summarized. In Sect. 5 the results are shown and finally in the last section, conclusions are commented.

## 2 Velocity Estimation

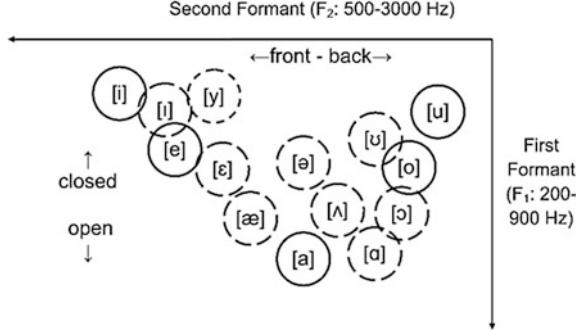
In speech production, the resonating cavities modifiable by the articulatory organs allow the energy of the speech signal to be focused at certain frequencies (formants), due to oropharyngeal tract resonators. The formant estimation is obtained from a digital inverse filter formulation. Iterative Adaptive Inverse Filtering algorithm [22], can give an adequate estimation of the glottal excitation. Moreover, a linear prediction model based on an autoregressive process (AR) [23] is enough to determine the formants (in non-nasal phonation) [24]. The first (F1) and second (F2) formants correspond to the two first maximum values in the linear prediction (LPC) spectrum.

In the vowel phonation the formants F1 and F2 vary for each vowel creating a vowel triangle [2] (Fig. 1).

The first formant is due to the muscles involved in jaw movement and the second formant due to tongue movement. These movements can be correlated with the formants positions in the plane F1 versus F2 [9] as:

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} \Delta F1 \\ \Delta F2 \end{bmatrix} \quad (1)$$

**Fig. 1** Vowel representation spaces adapted from Spanish (full circle) and American English (long-dash circle)



where  $\Delta x$  and  $\Delta y$  are the relative displacement from the previous position of the jaw and tongue.  $c_{ij}$  are the weights of the combination matrix.

Once the displacement is calculated, the velocity signal  $\vec{v}(t)$  is estimated as:

$$|\vec{v}(t)| = \frac{\sqrt{\Delta x(t)^2 + \Delta y(t)^2}}{\Delta t} \quad (2)$$

### 3 Sigma-Lognormal Model Application

#### 3.1 Overview

In this work, Sigma-Lognormal model [25] is used to parameterize the velocity profile of the tongue movement. This model considers the resulting velocity of a neuromuscular system action describing a lognormal function ( $\wedge(t)$ ) [26]:

$$D \wedge (t; t_0, \mu, \sigma) = \frac{D}{\sqrt{2\pi}(t - t_0)} e^{\left(\frac{-(\ln(t-t_0)-\mu)^2}{2\sigma^2}\right)} \quad (3)$$

where D is the scaling factor and  $t_0$  the time occurrence of the command.

Summing up each resulting lognormal vector, the complex movement pattern is given by equation:

$$\vec{v}_n(t) = \sum_{i=1}^M \vec{v}_j(t) \quad (4)$$

where  $\vec{v}_j(t)$  is the velocity profile of the  $j$ th stroke (simple movements involved in the generation of a given pattern of each neuromuscular impulse) and M represents the number of strokes.

The error between the original and its reconstructed signal gives the reconstruction quality in the sigma-lognormal domain. This could be evaluated using the Signal-to-Noise-Ratio (SNR) between the reconstructed velocity profile ( $\vec{v}_n(t)$ ) and the original one ( $\vec{v}(t)$ ). The SNR is defined as [27]:

$$SNR = 10 \log \left( \frac{\int_{t_s}^{t_n} [\vec{v}(t)^2] dt}{\int_{t_s}^{t_n} [(\vec{v}_n(t) - \vec{v}(t))^2] dt} \right) \quad (5)$$

### 3.2 Parameters

The proposed parameters from the reconstructed sigma-lognormal velocity profiles to measure the speech quality are the following:

- $\overline{\Delta t_0}$ : This parameter is the mean of the difference between the starting time of M consecutive strokes:

$$\overline{\Delta t_0} = \frac{\sum_{i=1}^M |t_{0i} - t_{0(i-1)}|}{M} \quad (6)$$

to measure the beginning time of each lognormal that conformed the velocity response.

- Mean of  $\mu$ : mean of the difference between the logtime delay ( $\mu$ ):

$$\overline{\Delta \mu} = \frac{\sum_{i=1}^M |\mu_i - \mu_{(i-1)}|}{M} \quad (7)$$

- Mean of  $\Delta \sigma$ : measure the differences between lognormal response time ( $\sigma$ ):

$$\overline{\Delta \sigma} = \frac{\sum_{i=1}^M |\sigma_i - \sigma_{(i-1)}|}{M} \quad (8)$$

- Nlog: Number of lognormals in 0.5 s in the central part of the phonation.
- SNR: Signal to noise ratio between reconstructed and original velocity profiles (Eq. 5).

## 4 Method

### 4.1 Subjects

Three different databases were used in the present study: two for the study of neurological pathologies (Parkinson and ALS) and the third for laryngeal diseases.

Firstly, a database including 3 Parkinson subjects was used from patients selected by neurologists. The subjects produced the five Spanish vowels. Each subject was asked to say the five Spanish vowels, in the same order and normal loudness and velocity. Each sample of the database comprises the five Spanish vowels (/a/, /e/, /i/, /o/, /u/ in the International Phonetic Alphabet) pronounced in a sustained way, lasting approximately 2 s for each vowel separated by silences.

Secondly, two cases of Amyotrophic lateral sclerosis (ALS) were recorded (in this case producing sentences in Portuguese). The target sentence was /tudo vale a pena quando a alma não é pequena/ (IPA: [tuðu̯ val̯e pInæ̯~kwæ̯n̯- duuał̯m̯e n̯a~ɛ~pk̯enæ̯~]) (All is worthwhile if the soul is not small). Control normal subjects were also recorded speaking the same sentence.

In the last two cases, all the samples were manually selected to avoid formant estimation artifacts.

Finally, a laryngeal disease and a healthy database were used to evaluate voice quality. For this purpose pathological (abnormal vocal fold behavior) and healthy voices are used. This database was recorded at the Hospital General “Doctor Negrín” in Gran Canaria (Spain) from vowels uttered by 20 pathological and 45 healthy subjects [16]. Half of the subjects (35) were male and the other were female. Each subject was asked to say the five Spanish vowels, in the same order and normal loudness and velocity, as in the first dataset. The pathological voices were obtained from speakers with a large range of speech system pathologies as hypo-function, vocal fold paralysis, polyps, etc.

### 4.2 Experimental Procedure

The experimental procedure was as follows:

First, the above mentioned parameters for subjects with neurological pathologies were compared with the healthy controls. The Anova analysis was used to measure the statistical significance of the above mentioned comparison. The compared results were considered different when the residual p-value was lower than 0.05 [28]. Bonferroni correction was not necessary as it is a simple comparison.

The above method was repeated for normal and laryngeal pathological subjects. The discriminative ability were worked out automatically using all the database.

## 5 Results

### 5.1 Normal Versus Parkinson

The voice was manually selected in the part of the signal where there was certainty that the formants were well estimated.

We can observe in Table 1, that there are two parameters showing a p-value lower than 0.05:  $\overline{\Delta t_0}$  and  $\overline{\Delta\sigma}$ . These two values are the lowest found in Parkinson speech. Figure 2 shows the results with  $\overline{\Delta\sigma}$ .

A small  $\overline{\Delta\sigma}$  in Parkinson speech may mean that Parkinson's Disease has limited the spread of neuromotor strokes, so the variation in the resulting movement is smaller than for a healthy subject.

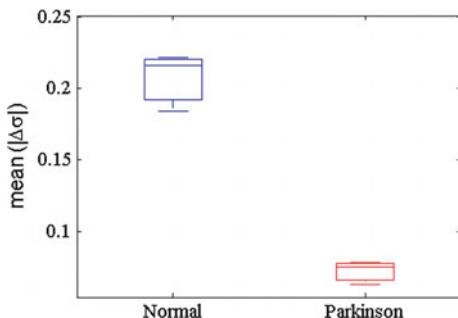
### 5.2 Normal Versus ALS

We also repeated the experiment with speech from two ALS patients. This experiment shows (Table 2 and Fig. 3) how the differences in the  $\overline{\Delta\sigma}$  parameter are the opposite than for the Parkinson case. We can intuitively think that the lognormals in the ALS case are wider than in the normal subjects but it is necessary to confirm this conjecture with more data.

**Table 1** Average parameters: normal versus Parkinson voices

Parameter	Normal	Parkinson	p-value
$\overline{\Delta t_0}$	<b>0.42</b>	<b>0.31</b>	<b>0.04</b>
$\overline{\Delta\mu}$	-1.3	-1.3	0.6
$\overline{\Delta\sigma}$	<b>0.22</b>	<b>0.07</b>	<b>&lt;0.01</b>
$Nlog$	14.15	12.33	0.1
$SNR$	24.04	20	0.25

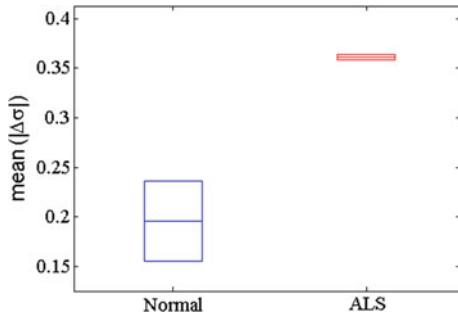
**Fig. 2** Parkinson versus normal phonation



**Table 2** Average parameters: normal versus ALS voices

Parameter	Normal	ALS	p-value
$\overline{\Delta t_0}$	0.40	0.34	0.624
$\overline{\Delta \mu}$	-1.13	-1.01	0.8
$\overline{\Delta \sigma}$	<b>0.19</b>	<b>0.37</b>	<b>0.045</b>
$Nlog$	<b>19.5</b>	<b>33</b>	<b>0.05</b>
$SNR$	24.46	23.35	0.68

**Fig. 3** ALS versus normal phonation



### 5.3 Normal Versus Laryngeal Pathology

We can observe in Table 3, that there are four parameters where it is possible to appreciate a p-value lower than 0.05:  $\overline{\Delta t_0}$ , NLog, SNR and  $\overline{\Delta \mu}$ . These four values are the lowest p-value found for laryngeal pathology speech.

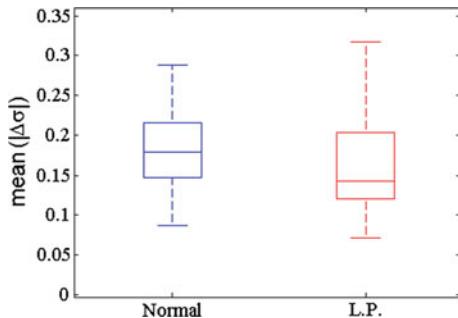
The comparison was repeated using the parameters for the five vowels and it shows that these parameters are vowel and gender independent.

Figure 4 shows how the  $\overline{\Delta \sigma}$  parameter is distributed in normal and laryngeal pathology speech and there is not statistically significant variation between healthy speakers and speakers with laryngeal pathology. A robust formant estimation is necessary to the correct kinematic estimation.

**Table 3** Average parameters: healthy versus laryngeal pathology voices

Parameter	Normal	Laryngeal pathology	p-value
$\overline{\Delta t_0}$	<b>0.36</b>	<b>0.286</b>	<b>&lt;0.01</b>
$\overline{\Delta \mu}$	<b>-1.19</b>	<b>-1.26</b>	<b>&lt;0.01</b>
$\overline{\Delta \sigma}$	0.18	0.14	0.88
$Nlog$	<b>10.90</b>	<b>8.87</b>	<b>&lt;0.01</b>
$SNR$	<b>24.12</b>	<b>22.07</b>	<b>&lt;0.01</b>

**Fig. 4** Healthy versus laryngeal pathology voices



**Table 4** Healthy versus pathologies voices

Parameter	Normal	Laryngeal pathologies	Parkinson	ALS
$\overline{\Delta\sigma}$	0.18	0.17	0.07	0.27

### 5.4 Discussion

Laryngeal pathologies are characterized by excitation problems (instability in the vibration frequency of the vocal cords, air flow and presence of noise). These diseases have no relationship with deterioration in the articulatory model.

In light of the results (Table 4), it seems that  $\overline{\Delta\sigma}$  do allow for characterization of differences in articulatory movement behavior, while  $\overline{\Delta\sigma}$  are no statistically significant variation between healthy speakers and laryngeal pathology speakers.

## 6 Conclusions and Future Study

In this work, the Sigma Lognormal model parameters have been calculated for different kinds of pathological speech. We have found that it is possible to get significant differences between healthy and laryngeal pathological speech with the parameters Nlog,  $\Delta t_0$  and the SNR. These parameters can be used as an indicator of the quality of the speech recording framework or the level of noise in speech.

Also we found that the discriminative ability of these features is independent of gender and vowel in our datasets.

As far as the case of neurological diseases is concerned, it has been found that the most relevant parameter is  $\overline{\Delta\sigma}$ .

Experimental results confirm that the application of the Sigma-Lognormal model to speech is a promising method to discriminate between healthy and different kinds of pathological speech.

The study has to be extended to more patients with Parkinson and ALS. Also, we could study how to improve formant extraction and how to improve the quality and robustness of formant features based on the Sigma-Lognormal parameters.

**Acknowledgments** This study was funded by the Spanish government's MCINN TEC2012-38630-C04 research project, the fellowship program of Universidad de Las Palmas de Gran Canaria and the NSERC-Canada Grant RGPIN-2015-06409 to R. Plamondon.

## References

1. Plamondon, R.: A kinematic theory of rapid human movements. Part I: Mov. Represent. Gener. Biol. Cybern. **72**, 295–307 (1995)
2. Plamondon, R., Djouia, M., Mathieu, P.: A time-dependence between upper arm muscles activity during rapid movements: observation of the proportional effects predicted by the kinematic theory. Hum. Mov. Sci. **32**, 1026–1039 (2013)
3. Gomez-Barrero, M., Galbally, J., Plamondon, R., Fierrez, J., Ortega-Garcia, J.: Variations of handwritten signatures with time: a sigma-lognormal analysis. In: Proceedings 6th International Conference on Biometrics, pp. 3.16.1–3.16.6 , June 4–7, Madrid, Spain (2013)
4. Plamondon, R., O'Reilly, C., Rémi, C., Duval, R.C.: The lognormal handwriter: learning, performing and declining. Front. Psychol. Cogn. Sci. Spec. Issue Cogn. Sci. 1–14 (2013)
5. O'Reilly, C., Plamondon, R.: Design of a neuromuscular disorders diagnostic system using human movement analysis. In: 11th International Conference on Information Sciences, Signal Processing and their Applications, Montreal, Canada, 3–5 July 2012
6. Plamondon, R., O'Reilly, C., Ouellet-Plamondon, R.: Strokes against strokes. Strokes Strides. Pattern Recogn. **47**, 929–944 (2014)
7. Carmona-Duarte, C., Alonso, J.B., Diaz-Cabrera, M., Ferrer, M., Gomez-Vilda, P., Plamondon, R.: Kinematic modelling of diphthong articulation. In: Proceedings of NOLISP 2015, volume of Smart Innovation Systems and Technologies, Springer (2015)
8. Peterson, G.E., Barney, H.L.: Control methods used in a study of the vowels. J. Acoust. Soc. Am. **24**–**2**, 175–184 (1952)
9. Gómez-Vilda, P., Londral, A.R.M., Rodellar-Biarge, V., Ferrández-Vicente, J.M.: Mamede de Carvalho: monitoring amyotrophic lateral sclerosis by biomechanical modeling of speech production. Neurocomputing **151**, 130–138 (2014)
10. Henríquez, P., Alonso, J.B., Ferrer, M.A., Travieso, C.M., Godino-Llorente, J.I., Díaz-de-María, F.: Characterization of healthy and pathological voice through measures based on nonlinear dynamics. IEEE Trans. Audio Speech Lang. Process. **17**(6), 1186–1195 (2009)
11. Alonso, J.B., de León, J., Alonso, I., Ferrer, M.A.: Automatic detection of pathologies in the voice by hos based parameters. EURASIP J. Appl. Signal Process. **2001**, 275–284 (2001)
12. Bandini, A., Giovannelli, F., Orlandi, S., Barbagallo, S., Cincotta, M., Vanni, P., Chiaramonti, R., Borgheresi, A., Zaccara, G., Manfredi, C.: Automatic identification of dysprosody in idiopathic Parkinson's disease. Biomed. Signal Process. Control 1747–54 (2015)
13. Tykalova, T., Rusz, J., Cmejla, R., Ruzickova, H., Ruzicka, E.: Acoustic investigation of stress patterns in Parkinson's disease. J. Voice **28**, 129 (2014)
14. Ramig, L.O., Fox, C., Sapir, S.: Speech treatment for Parkinson's disease. Expert Rev. Neurother. **8**, 297–309 (2008)
15. Galaz, Z., Mekyska, J., Mzourek, Z., Smekal, Zd., Rektorova, I., Eliasova, I., Kostalova, M., Mrackova, M., Berankova, D.: Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. Comput. Methods Prog. Biomed. (2016)
16. Skodda, S., Schlegel, U.: Speech rate and rhythm in Parkinson's disease. Mov. Disord. **23**, 985–992 (2008)
17. Rusz, J., Cmejla, R., Ruzickova, H., Ruzicka, E.: Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. J. Acoust. Soc. Am. **129**, 350–367 (2011)

18. Rusz, J., Cmejla, R., Ruzickova, H., Klempir, J., Majerova, V., Picmausova, J., Roth, J., Ruzicka, E.: Acoustic assessment of voice and speech disorders in Parkinson's disease through quick vocal test. *Mov. Disord.* **26**, 1951–1952 (2011)
19. Ball, L.J., Beukelman, D.R., Pattee, G.L.: Timing of speech deterioration in people with Amyotrophic lateral sclerosis. *J. Med. Speech Lang. Pathol.* **10**(4), 231–235 (2002)
20. Weismer, G., Martin, R., Kent, R.D., Kent, J.F.: Formant trajectory characteristics of Males with amyotrophic lateral sclerosis. *J. Acoust. Soc. Am.* **91**, 1085–1098 (1992)
21. Yunusova, Y.: Articulatory Movements During Vowels in Speakers With Dysarthria and Healthy Controls. *J. Speech Lang. Hear. Res.* **51**, 596–611 (2008)
22. Alku, P.: Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Commun.* **11**(2), 109–118 (1992)
23. Itakura, F.: Line spectrum representation of linear predictor coefficients of speech signals. *J. Acoust. Soc. Am.* **57**(S1), S35–S35 (1975)
24. Xia, K., Espy-Wilson, C.Y.: October: a new strategy of formant tracking based on dynamic programming. In: Interspeech, pp. 55–58 (2000)
25. Plamondon, R., Djouia, M.: A multi-level representation paradigm for handwriting stroke generation. *Hum. Mov. Sci.* **25**, 586–607 (2006)
26. O'Reilly, C., Plamondon, R.: Development of a sigma-lognormal representation for on-line signatures. *Pattern Recogn.* **42**, 3324–3327 (2009)
27. O'Reilly, C., Plamondon, R.: Design of a neuromuscular disorders diagnostic system using human movement analysis. In: Proceedings of the 11th International Conference on Information Sciences, Signal Processing and their Applications, Montréal, pp. 787–792, Canada (2012)
28. Hogg, R.V., Ledolter, J.: Engineering Statistics. MacMillan, New York (1987)

# Exploring and Comparing Machine Learning Approaches for Predicting Mood Over Time

**Ward van Breda, Johnno Pastor, Mark Hoogendoorn, Jeroen Ruwaard,  
Joost Asselbergs and Heleen Riper**

**Abstract** Mental health related problems are responsible for great sorrow for patients and social surrounding involved. The costs for society are estimated to be 2.5 trillion dollar worldwide. More detailed data about the mental states and behaviour is becoming available due to technological developments, e.g. using Ecological Momentary Assessments. Unfortunately this wealth of data is not utilized: data-driven predictive models for short-term developments could contribute to more personalized interventions, but are rarely seen. In this paper we study how modern machine learning techniques can contribute to better models for predicting short-term mood in the context of depression. The models are based on data obtained from an experiment among 27 participants. During the study frequent mood assessments were performed and usage and sensor data of the mobile phone was recorded. Results show that much can be improved before fine-grained mood prediction is useful within E-health applications. Subsequently important next steps are identified.

---

W. van Breda (✉) · J. Pastor · M. Hoogendoorn  
Department of Computer Science, VU University Amsterdam,  
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands  
e-mail: w.r.j.van.breda@vu.nl

J. Pastor  
e-mail: j.pastor@student.vu.nl

M. Hoogendoorn  
e-mail: m.hoogendoorn@vu.nl

J. Ruwaard · J. Asselbergs · H. Riper  
Department of Clinical Psychology, VU University Amsterdam,  
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands  
e-mail: jruwaard@me.com

J. Asselbergs  
e-mail: j.a.g.j.asselbergs@vu.nl

H. Riper  
e-mail: h.riper@vu.nl

## 1 Introduction

Mental health problems have a high impact on the lives of patients, their social surrounding, and the society in general. It obstructs patients to learn, work or participate in society. Many affected therefore turn to professional help. The costs for society have been estimated at a startling amount of 2.5 trillion dollars per year worldwide in 2015, and will rise to 6 trillion dollars per year by 2030 [1]. These problems are the driving force behind health-related research that aims to provide more effective therapies for patients and their social environment. Major developments in mobile technologies offer new possibilities for mental health interventions, such as performing Ecological Momentary Assessments (EMA) more effectively. Using new technology you can more frequently assess the mental state of the user as well as the context in which it is measured. The context can be collected unobtrusively using sensors. Such measurements provide highly detailed insights into the behavior and mental state of the patient and can be a driver for more personalized and effective therapies.

Although studies that involve EMA are increasing in number, only very few studies try to fully take advantage of the wealth of data that results. Predictive modeling on a more detailed level, e.g. predicting mood level changes in terms of hours, are rarely seen while they can be a great driver for more real-time (semi) automated forms of therapy. Most predictive modeling endeavors focus on more long term predictions such as therapeutic effectiveness or long term recovery. There are some exceptions, such as [2] or [3]. However, these studies do not take advantage of more recent developments in the area of machine learning that can lead to more accurate results and as a consequence show relatively poor performance.

In this paper, we try to utilize sophisticated machine learning techniques to accurately predict the mood within the context of depression. Depression has a life-time prevalence of 17.1 %, making it a major health problem [4], associated with morbidity, mortality, disability and psychological agony for the sufferers and their social surrounding [5]. Our starting point is a dataset collected among 33 participants in which sensory information from the smart phone of the participants has been collected as well as regular self-ratings of the mood. The current focus for the modeling is on predicting the current mood state of individual participants based on their individual histories as well as their current measurements using the smart phone. From the raw dataset, we derive a set of attributes following [2]. We then explore several ways to predict the aforementioned mood, specifically by using time-series techniques, dynamic time warping techniques, and a number of state-of-the-art machine learning techniques to see whether better techniques can contribute to more accurate predictions.

This paper is organized as follows. First, specific information about the data we will be using is discussed in Sect. 2, followed by the techniques we apply in Sect. 3. The results of our experiments are described in Sect. 4, and a discussion about the results can be found in Sect. 5.

## 2 Data

The data used in this research paper originates from the VU Unobtrusive Ecological Momentary Assessment pilot study data (for more info see [3]). First, the setup of the pilot study will be discussed, followed by the precise data that has been collected.

### 2.1 Pilot Study Setup

In the pilot study, measurements were performed over a period of approximately six weeks. A total of 33 participants were selected for the pilot of which 27 contributed enough data for meaningful analysis. The data consists of 76 variables and 1249 observations, running for 52 days. The data was obtained through two applications installed on the participant's smartphone: the eMate EMA application and the iYouVU application. The eMate EMA application prompts the user to rate their mood five times per day (at 09:00, 12:00, 15:00, 18:00 and 21:00) on a unidimensional scale as "mood" on a scale from 1–10, and the iYouVU application is a sensor logger, which is active in the background, unnoticeable to the participant. The duration for which the participants logged data varies and manual input was not always provided. In case the participants expressed their mood multiple times a day, an average was calculated for each day.

### 2.2 Data Description

For each participant, for each day an averaged mood value is present, which is referred to as an obtrusive feature attribute, because the participant manually needs to input their mood level over the day. The variables obtained through the iYouVU application are referred to as unobtrusive, because they are measured automatically. The values are aggregated per day. The number of calls and SMS messages to the top 5 contacts are measured as normalized frequency values. The duration of calls made to the top 5 contacts was measured, and the frequency and duration of the applications are measured as well. The appCat.n and appCat.duration are attributes representing the number of uses and duration for application categories. These variables are normalized per day between all categories of their respective class. The categories include: android, books, browser, business, education, entertainment, game, life style, email, music, news, productivity, social, tools, transportation, and unknown. The following attributes are also included: the number of images taken (image.n), the average screen duration per screen-on moment (screen.duration) and the screen-on frequency (screen.n), which were all normalized within the participant. Another variable obtained through the iYouVU application is the average percentage of accelerometer data that is classified as "high" (accelerometer.high). A summary of the dataset attributes are shown in Table 1.

**Table 1** The attributes that are present in the dataset in the form of averages per day

Attribute name	Type	Information	Range
mood	T	Daily mean unidimensional to be predicted mood	[0, 1]
mood.11	OF	Current mood, daily average	[0, 1]
call.c1c - call.c5c	UF	Number of calls made to top call contact 1–5	[0, 1]
call.c1d - call.c5d	UF	Duration of calls made to top call contact 1–5	[0, 1]
sms.c1c - sms.c5c	UF	Number of SMS sent to top SMS contact 1–5	[0, 1]
app.a1c - app.a5c	UF	Number of times top app 1–5 was launched	[0, 1]
app.a1d - app.a5d	UF	Duration of use of top app 1–5	[0, 1]
appCat.n	UF	App use frequency for each app category	[0, 1]
appCat.sum	UF	App use duration for each app category	[0, 1]
screen.duration	UF	Mean screen-on moment (standardized)	[0, 1]
screen.n	UF	Screen-on frequency (standardized)	[0, 1]
image.n	UF	Number of photos taken on smartphone	[0, 1]
accelerometer.high	UF	Mean percentage of accelerometer.high data	[0, 1]

Most of the where T is the target feature, OF is an obtrusive feature, and UF is an unobtrusive feature

### 3 Method

In this section we describe the methodology we have followed for generating our predictive models. We begin by discussing the steps taken to prepare the data: the process of derivation of additional attributes on top of the attributes described in Sect. 2 and the imputation of missing values. Then, we describe the techniques deployed on the dataset.

#### 3.1 Data Preprocessing

**Attribute Engineering** The basic attributes which have directly been based on measurements in the pilot have already been discussed in Sect. 2. While these attributes are certainly useful, new information that can be derived from these attributes could potentially also be beneficial. We therefore constructed a number of additional attributes, that we included in a separate analysis. One attribute that was added concerns the weekday, which was calculated by using the already available data/time attribute. Another two attributes were the sum of call request and/or SMS to the top 5 persons for each day, and the sum of application usage of applications belonging to the top 5 used applications each day.

**Missing value imputation** The dataset contains several missing values for both the measurements of the smart-phone sensors as well as the EMA data. We consider data

as missing if there is no single measurement on a day (the granularity of measurements considered in this case). There are a number of ways to deal with missing data; time points with missing values can either be removed or the missing values can be imputed. As already mentioned in the data description, the number of measurement points is quite limited. Therefore, the decision was made to only remove the time points of which the variable mood was missing, and impute the other values.

There are a total of 1249 time points, where for 1224 the mood variable is available. 1099 observations are without any missing values. We chose to impute the missing values with the mean of the variable per participant, using all observations of that individual.

### 3.2 Predictive Models

Different methods can be used to make numerical predictions. The decision for which predictive models to use in this experiment was made based on their properties. Predictive models we considered preferably are state-of-the-art algorithms, are known to perform well, can be used for regression, do not require much parameter tuning and do not take an excessive amount of time to train models. On a higher level the rationale behind selecting these predictive models for comparison is that we want to apply techniques that (1) only consider trends in the target based on its previous value (i.e. predict future mood based on past developments of mood): *mood only*, (2) try to use previously seen participants and predict based on their data: *similarity*, and (3) that use the feature attributes and prior values of mood for predicting future mood: *full predictive modeling*. Each of the three options mentioned above are explained in more detail below.<sup>1</sup>

**Mood only** For predicting mood exclusively using the mood series for each individual, we start with time series modeling using Auto Regressive Integrated Moving Average (ARIMA). The ARIMA model explores if the mood signal is a ‘stationary’ signal, i.e. the model tries to find statistical properties that make the mood series constant over time.<sup>2</sup>

**Similarity** To explore similarities in the mood series the Dynamic Time Warping (DTW) algorithm was applied. This algorithm compares two time-series for similarities, even if they differ in terms of time or speed. Such similarities could possibly exist between different individuals related to their mood series, and then possibly be used for predicting mood of individuals with highly similar mood series patterns. Since the DTW algorithm cannot deal with missing values, these values were

<sup>1</sup>All models were constructed making use of R [6].

<sup>2</sup>For fitting the ARIMA model the Forecast package [7] was used (AICc with a correction for finite sample sizes).

imputed using the average of the variable before and after the missing value. This imputed sequence was normalized after which the algorithm was applied.<sup>3</sup>

**Full predictive modeling** For full predictive modeling we use two techniques: Support Vector Machines (SVM) and Random Forests (RF). SVM generates hyperplanes in a multidimensional space for purpose of classification or regression. Our implementation of SVM uses a gaussian radial basis function for the regression task.<sup>4</sup> The RF is a technique that averages the performance of a number of decision trees. The individual decision trees are trained on different parts of the dataset, and together do not suffer from overfitting problems.<sup>5</sup>

### 3.3 Setup and Measurements

**Mood only** For each participant the ARIMA models use the whole series as training data. This first step, i.e. to reproduce the data, is important because we want to know the potential power of the method. Thus, the time series model has maximum potential to find statistical attributes to explain the dynamics for each individual. After this, for each individual the tuned model is fed the data related to each time point in an attempt to predict the target variable (the mood value) for each next time point. The final MSE for this method is calculated by taking into account the MSE for all individuals and all time points.

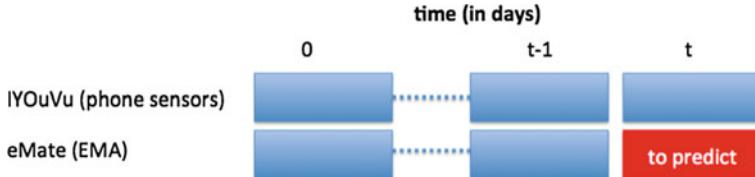
**Similarity** The similarity measure uses the whole series as training data as well, because we want to see the extent of similarities that are present between participants in the most ideal circumstance. In such a way we can better assess the potential of similarity measures for predicting mood series. Regarding the evaluation setup however, as is discussed in Sect. 4, we do not believe this has much potential yet. Therefore no evaluation setup is necessary for this method. Related to the settings of the algorithm, the “Sakoe-Chiba” band was used with a maximum window size of 6. The hierarchical clustering that was applied used the complete-linkage clustering method.

**Full predictive modeling** For each participant the SVM and RF methods are trained based on the past days that are available and predict the target variable, the mood value, for the next time point. This way, as time increases, more training cycles become available for the machine learning method to train on. So, starting on day three, each model has two training cycles, predicting the mood for the current day; on day four, each model has three training cycles, and so on. The final MSE for each method is calculated by taking the MSE for all individuals for all time points. Due to

<sup>3</sup>For using this technique the DWT package [8] was used, supported by [9] to allow for open begin/end comparisons. The clustering of time series was done making use of the TSClust package [10].

<sup>4</sup>For implementation we used the kernlab package [11].

<sup>5</sup>For implementation we used the random Forest package [12].



**Fig. 1** Overview learning setup

this setup, the methods tend to have better performance as more training data cycles become available over time. The configuration for learning the algorithms is shown in Fig. 1.

Concerning the settings of the methods; for the SVM method, the epsilon regression was applied using a radial kernel, with the cost of constraint violation set to .5, epsilon to .1 and the sigma was determined by the built-in hyperparameter estimation heuristic function signet [13]; and for the RF method, the number of variables randomly sampled as candidates at each split was set to the number of columns divided by 3, the number of trees to grow to 500. Because fitting 1170 models (without parameter tuning) takes a considerable amount of time, no further parameter tuning on a per model basis was done as of yet, as this would further increase the computation times.

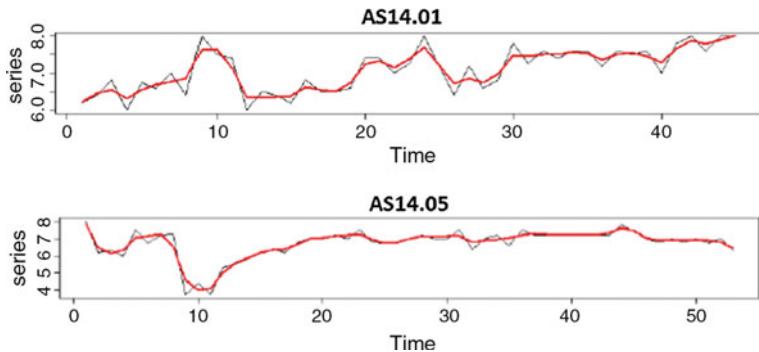
## 4 Results

As discussed in Sect. 3, we applied three types of methods: one where only the mood is considered for predicting mood for the next time point, one where we explore similarity analysis methods for predicting mood, and one where we apply full predictive modeling methods, namely RF and SVM. To compare the performance of the methods we added a benchmark method which simply predicts the same value as the average mood of all available past days.

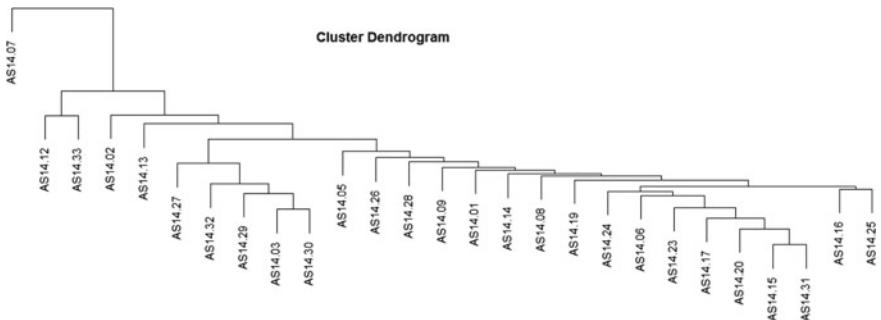
### 4.1 Mood Only

We start with time series analysis by deconstructing the univariate mood series of participants by representing them as ARIMA series. In this case we started with exploring how well the ARIMA is able to reproduce the trends seen for the mood as this is critical as a first step towards prediction. For exemplary purposes in Fig. 2 the fit of the ARIMA models on the mood series of participant 1 and 5 is shown.

Based on being trained on beforehand on the full mood series for each individual. The ARIMA model produces an MSE of 0.475 for the reproduction of the mood



**Fig. 2** Participant 1 and 5—best fit of the ARIMA models on the participant mood series



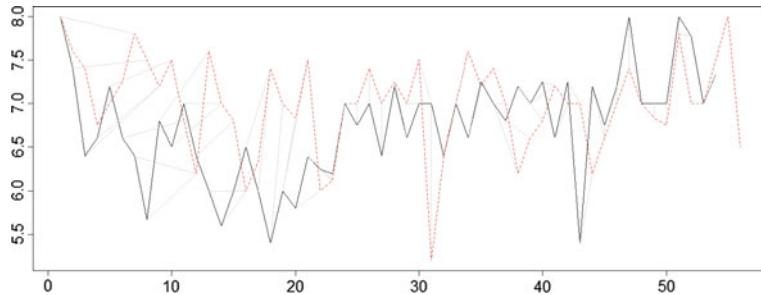
**Fig. 3** Dendrogram of clustered time series

patterns. Since this is even below the naive benchmark explained before (0.442) we conclude that using mood only is not enough to predict mood series accurately for individuals.

## 4.2 Similarity

Next, we applied Dynamic Time Warping to see whether patients exhibit similar patterns and could potentially be used to predict the mood of unseen patients. The application of this approach resulted in a matrix with associated similarities between the different participants. We applied hierarchical clustering on this matrix.<sup>6</sup> The resulting dendrogram can be found in Fig. 3. It may be concluded that the mood series of participant 7 is quite unlike that of the other participants. Furthermore the dendrogram provides insight into how the participants are clustered and provides a quick view into how a participant compares to others participants. An example of two

<sup>6</sup>Specifically referred as UPGMA.



**Fig. 4** Similarity DTW participants 16 and 24

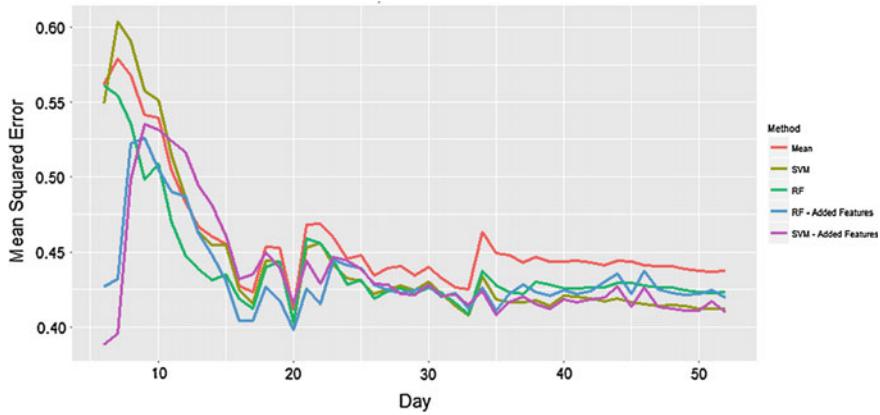
similar time series and the resulting DWT path is illustrated in Fig. 4, where the series of participant 16 and 24 are plotted. Although being an interesting analysis, as can be seen in Fig. 4, based on the great variation in patterns and limited number of participants we did not consider this a viable option for prediction either.

### 4.3 Full Predictive Modeling

Finally, we applied the more traditional machine learning approaches. Here, we varied whether we used the added features (cf. Sect. 2 or not. Without these features, the SVM model obtains an MSE of 0.411, and the RF models scores 0.425. With the added features, the SVM models perform with a MSE of 0.410, and the RF models with a MSE of 0.420. In any case both methods are able to outperform the naive benchmark method, with an MSE of 0.442. Both SVM and RF did have a slight advantage with the added features. An overview of the performance of each of the methods is displayed in Table 2. Figure 5 shows how the MSE changes as more historical data is provided to the algorithm. It can be seen that predictive performance stabilizes after around 20 days of historical data. Note that the MSE is only measured for the day following the number of days considered as history.

**Table 2** The performance of the different algorithms

Method name	MSE
Benchmark method (Naive)	0.441
SVM without added features	0.411
RF without added features	0.425
SVM with added features	0.410
RF with added features	0.420



**Fig. 5** The cumulative performance (MSE) over time for all participants. The x-axis shows the number of days of history considered

## 5 Discussion

There is limited work on using sensor data to predict short-term developments of mood and related aspects, especially studies that have used machine learning approaches for this purpose. In this paper, we have therefore taken a more exploratory approach and applied a number of well-known machine learning techniques to a dataset collected in a small pilot study. We looked at the possibility to use time-series models to predict mood, where only the mood series itself was used without the measurements using the sensors on the phone. Thereafter we did include the unobtrusive measurements combined with dynamic time warping and more classical machine learning approaches.

Predicting mood based on past mood alone turned out to have low predictive performance. Mood generally does not seem to have intrinsic statistical properties that can explain the dynamics of mood to a large degree. Also, looking at similarities between mood series of different individuals, it was found that certain pairs of individuals do share more similarities than others, but we found the similarities too limited to be of use for prediction. Finally, the predictive models using SVM and RF that leveraged all data did result in better performance compared to the naive benchmark and the time-series models. This finding indicates that contextual data about the individual is useful and needed to increase predictive performance, but, given the attributes we had to our disposal, is not yet at the level to meaningfully employ within an E-health application. We think that the focus should be on finding the most relevant attributes that highly correlate with the dynamics of the target variable (in our case short-term mood), and think about how to optimally measure such attributes. Only then a sharp increase in predictive performance can be expected.

Next to finding new meaningful variables related to the target variable, we think preprocessing the data is an very important part of the process. The feature attributes we added in the preprocessing stage enabled a small increase in performance but we feel there is still a lot to gain.

## References

1. Insel, T.R., Collins, P.Y., Hyman, S.E.: Darkness invisible: The hidden global costs of mental illness. *Foreign Aff.* **94**, 127 (2015)
2. LiKamWa, R., Liu, Y., Lane, N.D., Zhong, L.: Moodscope: building a mood sensor from smartphone usage patterns. In: Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys '13, pp. 389–402. ACM, New York (2013)
3. Asselbergs, J., Ruwaard, J., Ejdys, M., Schrader, N., Sijbrandij, M., Riper, H.: Smartphone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *J. Med. Internet Res.* (forthcoming)
4. Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H.-U., Kendler, K.S.: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the united states: results from the national comorbidity survey. *Arch. Gen. Psychiatry* **51**(1), 8–19 (1994)
5. Murray, C.J., Lopez, A.D.: Alternative projections of mortality and disability by cause 1990–2020: global burden of disease study. *Lancet* **349**(9064), 1498–1504 (1997)
6. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2014)
7. Hyndman, R.J., Khandakar, Y.: Automatic time series forecasting: the forecast package for R. *J. Stat. Softw.* **26**(3), 1–22 (2008)
8. Giorgino, T.: Computing and visualizing dynamic time warping alignments in R: the DTW package. *J. Stat. Softw.* **31**(7), 1–24 (2009)
9. Tormene, P., Giorgino, T., Quaglini, S., Stefanelli, M.: Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artif. Intell. Med.* **45**(1), 11–34 (2008)
10. Montero, P., Vilar, J.A.: TSclust: an R package for time series clustering. *J. Stat. Softw.* **62**(1), 1–43 (2014)
11. Zeileis, A., Hornik, K., Smola, A., Karatzoglou, A.: Kernlab—an S4 package for kernel methods in R. *J. Stat. Softw.* **11**(9), 1–20 (2004)
12. Liaw, A., Wiener, M.: Classification and regression by randomforest. *R News* **2**(3), 18–22 (2002)
13. Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A.: kernlab—an S4 package for kernel methods in R. *J. Stat. Softw.* **11**(9), 1–20 (2004)

# Cross-Cultural Telepsychiatry: An Innovative Approach to Assess and Treat Ethnic Minorities with Limited Language Proficiency

Davor Mucic

**Abstract** Current refugee crisis within European Union (EU) challenges mental health care systems in each EU country. For ethnic minorities in EU access to mental health care is a problem due lack of clinicians who understand their language, culture and special needs. Linguistic, cultural and even racial differences between patient and provider can have an impact on the therapeutic alliance. Therefore communication between providers (mental health professionals) and cross-cultural patients is even more complicated with a third person i.e. interpreter, involved. However, refugees and asylum seekers still receive the most of treatment provided via interpreters. Innovative solution for this problem might be “cross-cultural telepsychiatry model” within various settings. Since 2004, “cross-cultural telepsychiatry” has been tested, developed and established in outskirts areas of Denmark through various pilot projects. Overall high patient satisfaction was reported by patients as well as by involved professionals.

**Keywords** e-Mental health · Telepsychiatry · Asylum seekers · Refugees and migrants · Access to care · Language barriers · Interpreters · Cross-cultural (tele) psychiatry

## 1 Introduction

Accessibility and availability of culture-appropriate care are critical pathways to establishment of an effective care system. However, the access to relevant care is not always an easy task for e.g. asylum seekers, refugees and migrants within EU. Language and cultural barriers are among greatest obstacles for cross-cultural patient population to access adequate care. The patients with limited language proficiency are less likely to receive empathy, establish rapport, receive information

---

D. Mucic (✉)

Little Prince Treatment Centre, Copenhagen V, Denmark

e-mail: info@denlilleprins.org

URL: <http://www.denlilleprins.org>

and encouragement to participate in decision making [1–7]. Language barriers are associated with lower rates of patient satisfaction and poor care delivery in comparison with care received by patients who speak the language of the care provider [8, 9]. The patients who face language barriers are less likely than others to have a usual source of medical care; frequently receive preventive services at reduced rates; have an increased risk of non-adherence to medications; are less likely than others to return for follow-up appointments after visits to the emergency room; and have higher rates of hospitalization and drug complications [10].

The presence of a third person i.e. an interpreter, in a confidential relationship affects patient satisfaction, as it influences both transference and counter transference between individuals involved, with unavoidable consequences on a doctor-patient relationship [11]. Further, interpreter mediated communication is linked to increase risk of loss of confidentiality which is why the most asylum seekers, refugees and migrants exposed for such kind of communication tend to be suspicious wondering: “How soon will everyone in this little city speak about my illness?” [12]. It is not unusual as among ethnic and racial minorities, in comparison to the majority group, mental illness may be even more stigmatized [13]. Further, no matter whether the treatment is acute or long term, the consequences of interpreter provided care are increased treatment expenses and much longer time used to assess and/or treat each patient. Consequently, patient satisfaction, treatment alliance as well as compliance are affected.

On the other hand, adequate language concordance is significantly associated with higher reporting of past experience of traumatic events and of severe psychological symptoms, contrasting with much fewer referrals to psychological care when language concordance is inadequate [14]. Clearly, language- and even racial-concordance are associated with better patient compliance, better adherence to treatment, and higher patient satisfaction within mental health as well as in other health-care settings [15–17].

“Ethnic matching” appears to be the most desirable model used in addressing language barriers and cultural disparities in mental health-care provision of cross-cultural patient population. The term covers over the use of culturally competent bilingual clinicians who have the same ethnic and cultural background as their respective patients. Ethnic matching, supplemented by culture competency training, has been proved as a common strategy to address a number of barriers in cross-cultural related health-care provision [18, 19].

However, ethnic matching is not that easy to implement. When the patient and the ‘matching’ clinician are located in different places then a consultation is likely to require travel, either for the patient or the clinician. Effective responses to issues mentioned above require innovation, the capacity to “think outside the box”, culturally competence, and institutional support. One solution is to bring “cultural expertise” to the patient by use of the e-Mental Health (eMH). eMH is the use of telecommunication and information technologies to deliver mental health services at a distance [20]. eMH interventions have a number of advantages: They are easily accessible, provide anonymity to the user and are less expensive than patient-provider contacts in-person [21].

Telepsychiatry is the most described and evidence based form of eMH that enables patient and doctor to see and hear each other and interact regardless the distance. While various telepsychiatry applications have been tested and developed over the last five-six decades, there are relatively few published reports describing the use of telepsychiatry in the provision of mental health care to cross cultural patients [22–25]. Innovative model of telepsychiatry is “Cross Cultural Telepsychiatry” (CCT). It covers the delivery of culturally appropriate mental health care from a distance. It can be done in “real- time” by the use of videoconferencing (synchronous telepsychiatry) and more recently developed asynchronous telepsychiatry (“store and forward” model), where we speak about a transmission of recorded clinical related material i.e. assessment, psychiatric interview/consultation between referring physicians and specialist. The clinical service may include the interview, other assessment, psychiatric consultation between referring physicians and specialist, and other components [26].

## 2 “Cross-Cultural Telepsychiatry Model”

The first CCT pilot project in Denmark was developed in period 2004–2007. The aim of the project was to overcome the burden of poor service access for ethnic minorities in Denmark and promote a new way of delivering mental health care by use of videoconferencing in real-time [23]. Thereafter, different approaches have been described dealing with specific needs of Hispanics/Latinos and Asians [27–30] and Native American [31, 32].

The hypothesis behind the development of CCT in Denmark was that the majority of cross-cultural patients would prefer contact in their mother tongue, even when provided via telepsychiatry, rather than interpreter provided in-person contact with a Danish doctor.

Little Prince Treatment Centre in Copenhagen has telepsychiatry cross-cultural expertise more than other places in Europe [33]. The Centre is a private clinic specialized in treatment of ethnic minorities where affiliated clinicians are bilingual, cultural competent mental health professionals. Four stations (i.e. two hospitals, one asylum seekers centre and one social institution for rehabilitation of refugees and migrants) were connected via videoconference with Little Prince Treatment Centre in Copenhagen. Bilingual clinicians affiliated to the Centre assessed and/or treated patients via their own language, providing reliable assessment and valid treatment for a wide variety of psychiatric disorders. A patient satisfaction questionnaire was special designed (Table 1) for completion at the end of the visit. High acceptance and satisfaction regardless the patients’ ethnicity or educational level was reported [34]. All patients preferred “remote” contact compared to in-person care with a interpreter, due to perceived higher anonymity, confidence/trust in providers and self-efficacy to express intimate thoughts and feelings without a third person involved. As expected, there was a clear correlation between the number of sessions, reported satisfaction level, and quality of care [34].

**Table 1** Questionnaire

		Yes, in high degree (%)	Yes, in some degree (%)	No, only in less degree (%)	No, not at all (%)
1	Did you get enough information about telepsychiatry?				
2	Do you perceive "contact via TV" as uncomfortable?				
3	Did you feel safe under telepsychiatry contact?				
4	Have you been satisfied with sound quality?				
5	Have you been satisfied with picture quality?				
6	Did you achieve your goal via telepsychiatry/could you express everything you wanted to?				
7	Would you recommend telepsychiatry to others?				
8	Would you prefer contact via translator in future?				
9	What advantages did you perceive by telepsychiatry contact?				
10	What disadvantages did you perceive by telepsychiatry contact?				

A sustainable telepsychiatry service between Psychiatric department on island Bornholm and Little Prince Treatment Centre remained functioning after the first pilot project ended in 2007. It is to our knowledge the only such service in EU.

### 3 Telepsychiatry Within Hospital Setting

CCT assessments of hospitalized suicidal cross-cultural patients are particularly useful especially when it comes to patients that have had a telepsychiatric contact prior to involuntarily admission. Narratives from daily clinical work may significantly increase the understanding and acceptance of telepsychiatry among professionals with no telepsychiatry related experiences or professionals that are still in doubt. The following episode occurred within the first telepsychiatry pilot project:

*NN, 38 y.o. male, refugee from Bosnia-Herzegovina, diagnosed with PTSD and treated via telepsychiatry for 1 year prior to involuntary hospitalization caused by increased suicide risk and suicidal threats that NN presented for his general practitioner who decided to send NN to psychiatric emergency department located on the island where NN lives. There NN was assessed by Danish psychiatrist via Bosnian interpreter and involuntarily hospitalized. A day after NN was seen by the psychiatrist who treated NN via telepsychiatry prior to hospitalization. It was convenient for psychiatric department located on isolated island to call the psychiatrist that speak the same language as the patient in order to assess the patient's mental state, including the current risk of suicide. Despite the fact that the consultation has been done remotely NN could disclose much more via videoconference on mother tongue than via interpreter provided in-person consultation with Danish doctor the day before.*

## 4 Telepsychiatry Shared-Care Model

Involuntary admissions are relatively frequent among mentally ill ethnic minorities compared to domicile population. This might be due to poor contact with General Practitioners (GPs) and outpatient psychiatric services, and thus at greater risk of serious deterioration in mental illness before treatment is started [35–37]. Most often GPs, are not provided by equal access to psychiatric supervision and expertise. The necessity of sending patients onwards in the system often ends with long waiting time during which patients are usually not given any help, and their mental condition worsens. According to scientific research, there is increasing recognition that improving the detection, treatment and outcomes for mental health problems requires service models that integrate mental health care within primary health care practice [38]. “Shared care” covers a broad spectrum of collaborative treatment arrangements and there is no standard definition in the literature [39, 40]. Nevertheless, shared care successfully integrate and link mental health services with primary care [41]. When joint primary and specialist level collaborative care models have been evaluated using RCT designs, a range of clinical and service benefits are reported [42]. The Little Prince Psychiatric Centre was in charge of a project which offered an alternative approach by applying telepsychiatry provided shared care model i.e. collaborative care model via videoconference.

Little Prince Treatment Centre in cooperation with 6 GP clinics on the outskirts of Denmark conducted a shared-care pilot project in period July 2010-December 2015. Patients were both “domestic” (of Danish origin) and ethnic minorities. The results has shown that collaboration via use of videoconferencing across levels of health care sectors can be a useful alternative that offers learning, leads to continuity, reduces costs and improves the quality of care. Telepsychiatry has been well received by patients and general practitioners as a method reducing waiting time and bridging the distance between patients and specialized psychiatric care. GPs involved in the project perceived the service as a valuable and effective supplement to already existing practice.

## 5 Telepsychiatry in Vocational Rehabilitation

A study of the Danish labour market showed that mental illnesses and psychological difficulties are some of the main reasons impeding a growing number of unemployed individuals from entering and integrating into the labour market [43]. Further, active functioning in the labour market has been identified as one of the seven targets for successful integration of foreigners into the Danish society [44].

The Job Centres in Denmark are responsible for providing help in finding jobs and granting disability pensions. One of the specific responsibilities of a job center is to implement a vocational rehabilitation offer, which is usually relevant for individuals with physical, psychological and/or social difficulties. A vocational rehabilitation offer in Denmark means implementation of different activation programs such as education or job training and state financial support. In order to offer a vocational rehabilitation program or grant a disability pension, a caseworker at the job center has to explore and confirm, that a working capacity of the individual in question is reduced. In case of suspected mental illness, the clarification of vocational capacity requires a professional statement from a psychiatrist.

A total of ten job centres, located in different parts of Denmark, 64–299 km away from Little Prince Treatment Center in Copenhagen, have agreed to participate in the project. Twenty caseworkers were selected by the job centres. A three-phase pilot project was carried out in period July 2010–January 2012. Ten job centres, located in different parts of Denmark, participated by referring their clients to Little Prince Treatment Center. Mental health specialists with the relevant language skills conducted the assessment interviews via videoconference and generated an assessment report. A satisfaction questionnaire was completed by the caseworkers and the clients. Forty nine unemployed individuals were referred by caseworkers during a period of 19 months. A variety of psychiatric diagnoses was disclosed. The overall satisfaction with the telepsychiatry service was reported by both the clients and the caseworkers.

## 6 International Telepsychiatry

The most comprehensive international telepsychiatry service in the world was established in Denmark in mid-2006 (May 2006–October 2007) as a part of cross-cultural telepsychiatry pilot project mentioned above [45]. Because resources with cross-cultural skills were more readily available in Sweden than in Denmark, it was desirable to involve cross-cultural clinicians from Sweden. Videoconferencing equipment connected the Swedish Department of the Little Prince Psychiatric Centre with above mentioned 4 stations during period of 18 months. Overall, high patient satisfaction was reported and minor disadvantages of telepsychiatry were offset by the fact that the doctor-patient language and cultural matching acceptance. The use of bilingual

clinicians with a similar ethnic and cultural background to their patients compensates for the distance and lack of physical presence.

The crucial indicators of patient satisfaction were:

- Accessibility of culturally competent care via mother tongue;
- Ability to express intimate thoughts and feelings from a distance, without third person involvement;
- Perceived safety and comfort by the service;
- High quality of sound and picture;
- Time savings associated with no need for travel;
- Reported willingness to use telepsychiatry again and recommending it to others;
- Preference for telepsychiatry in comparison to interpreter-assisted care.

## 7 Discussion

Commonly, clinicians prioritize in-person contact with the patient. When the communication is preceded by long travel, long waiting time or involvement of an interpreter then both the clinicians and the patients may benefit by use of telecommunication technology. The large majority of the patients accepted the CCT model regardless the type of the service setting. The patients' judgement of enhanced safety and comfort by telepsychiatry might be due to less likelihood of meeting the doctor on the street and the risks of spreading rumours in the patients' neighbourhood; this is similar to Native American populations in small tribes in the U.S. [46]. Ethnicity did not appear to be associated with the patients' attitudes towards telepsychiatry. Differences in perception of the service were identified with respect to the patients' previous experiences with the mental health system in Denmark. Patients who had earlier received treatment via interpreters in Denmark were more favourable to mother tongue-provided telepsychiatry, compared to patients without previous interpreter-related experiences. It was easier for the patients to express themselves from a distance, and they actually felt more secure and could control the situation which resulted in them being more open.

However, there is no doubt that some patients will prefer remote consultations due to controlling the presence of the psychotherapist, so as to feel less influenced by the clinician i.e. having the opportunity to "switch off" the therapist.

Continuity remained by seeing the same doctor no matter where the patient is located is probably one of above described services most important advantages compared to traditional mental health care provision.

## 8 Conclusion

Even before current refugee crisis in EU, it has become increasingly evident that standard treatment approaches require modification or adaptation in order to ensure that cross-cultural patient population with limited language proficiency receive effective mental health care. The use of videoconference enables opportunity to build the bridges over cultural and linguistically barriers by connecting patients with professionals who “match” culturally and linguistically. Promising results of the first international telepsychiatry project, might pave the way for potential development of an international service where bilingual professionals all over the globe would be able to share their knowledge and expertise in order to assess and/or treat mentally ill ethnic minorities via respective mother tongue. Clinical and scientific objectives and goals of such international telepsychiatry service as well as the potential outcomes are endless [47]. Within clinical practice, we have never been presented with the tool that requires so little investment while it in turn gives us so much.

## References

1. Kaplan, S.H., Gandek, B., Greenfield, S., et al.: Patient and visit characteristics related to physicians' participatory decision-making style. Results from the medical outcomes study. *Med. Care* **33**, 1176–1187 (1995)
2. Cooper-Patrick, L., Gallo, J.J., Gonzales, J.J., et al.: Race, gender, and partnership in the patient-physician relationship. *JAMA* **282**(6), 583–589 (1999)
3. Cooper, L.A., Roter, D.L., Johnson, R.L., et al.: Patient-centered communication, ratings of care, and concordance of patient and physician race. *Ann. Intern. Med.* **139**, 907–915 (2003)
4. Sundquist, J., Winkleby, M.A.: Cardiovascular risk factors in Mexican American adults: a transcultural analysis of NHANES III, 1988–1994. *Am. J. Pub. Health* **89**, 723–730 (1999)
5. Saha, S., Arbelaez, J.J., Cooper, L.A.: Patient-physician relationships and racial disparities in the quality of health care. *Am. J. Pub. Health* **93**, 1713–1719 (2003)
6. Ferguson, W.J., Candib, L.M.: Culture, language and the doctor-patient relationship. *Fam. Med.* **34**(5), 353–361 (2002)
7. McGlade, M.S., Saha, S., Dahlstrom, M.E.: The Latina paradox: an opportunity for restructuring prenatal care delivery. *Am. J. Pub. Health* **94**, 2062–2065 (2004)
8. Carrasquillo, O., Orav, E.J., Brennan, T.A., Burstin, H.R.: Impact of language barriers on patient satisfaction in an emergency department. *J. Gen. Intern. Med.* **14**, 82–87 (1999)
9. Sarver, J., Baker, D.W.: Effect of language barriers on follow-up appointments after an emergency department visit. *J. Gen. Intern. Med.* **15**, 256–264 (2000)
10. Flores, G.: Language barriers to health care in the United States. *N. Engl. J. Med.* **355**, 229–231 (2006)
11. Spiegel, J.P.: Cultural aspects of transference and countertransference revisited. *J. Am. Acad. Psychoanal.* **4**, 447–467 (1976)
12. Mucic, D., Hilty, D.M., Yellowlees, P.M.: e-Mental Health Toward Cross-Cultural Populations Worldwide. In: Mucic, D., Hilty, D.M. (eds.) *e-Mental Health*, pp. 77–91. Springer, Cham (2016)
13. Leong, F.T.L., Kalibatseva, Z.: Cross-cultural barriers to mental health services in the United States. *Cerebrum: the Dana Forum on Brain. Science* **2011**, 5 (2011)

14. Bischoff, A., Bovier, P.A., Rrustemi, I., Gariazzo, F., Eytan, A., Loutan, L.: Language barriers between nurses and asylum seekers: their impact on symptom reporting and referral. *Soc. Sci. Med.* **57**(3), 503–512 (2003)
15. Saha, S., Komaromy, M., Koepsell, T.D., Bindman, A.B.: Patient-physician racial concordance and the perceived quality and use of health care. *Arch. Intern. Med.* **159**, 997–1004 (1999)
16. Bowen, S.: Language Barriers in Access to Health Care. *Health Canada*, Ottawa (2001)
17. Perez-Stable, E.J.: Language access and Latino health care disparities. *Med. Care* **45**, 1009–1011 (2007)
18. Ton, H., Koike, A., Hales, R.E., Johnson, J.A., Hilty, D.M.: A qualitative needs assessment for development of a cultural consultation service. *Transcult. Psychiatry* **42**, 491–504 (2005)
19. Jerrell, J.M.: Effect of ethnic matching of young clients and mental health staff. *Cult. Divers. Ment. Health* **4**, 297–302 (1998)
20. Mucic, D., Hilty, D.M. (eds.): *e-Mental Health*. Springer, Cham (2016)
21. Mock J (2014). Support from the Internet for individuals with mental disorders: advantages and disadvantages of e-mental health service delivery. *Front Public Health* 2:65, [10.3389/fpubh.2014.00065](https://doi.org/10.3389/fpubh.2014.00065)
22. Sherrill W.W., Crew, L., Mayo, R.M., et al.: Educational and health services innovation to improve care for rural Hispanic communities in the U.S. *Educ Health (Abingdon)* **18**(3), 356–367 (2005)
23. Mucic, D.: Telepsychiatry in Denmark: mental health care in rural and remote areas. *J. e-Health Technol. Appl.* **5**(3), (2007)
24. Yeung, A., et al.: A study of the effectiveness of telepsychiatry-based culturally sensitive collaborative treatment of depressed Chinese Americans. *BMC Psychiatry* **11**, 154 (2011)
25. Shore, J., Kaufmann, L.J., Brooks, E., et al.: Review of American Indian veteran telemental health. *Telemed. J. e-Health* **18**(2), 87–94 (2012)
26. Hilty, D.M., Ferrer, D.C., Parish, M.B., et al.: The effectiveness of telemental health: a 2013 review. *Telemed. J. e-Health* **19**, 444–454 (2013)
27. Nieves, J.E., Stack, K.M.: Hispanics and telepsychiatry. *Psychiatr. Serv.* **58**(6), 877 (2007)
28. Moreno, F.A., Chong, J., Dumbauld, J., et al.: Use of standard webcam and Internet equipment for telepsychiatry treatment of depression among underserved Hispanics. *Psychiatr. Serv.* **63**(12), 1213–1217 (2012)
29. Chong, J., Moreno, F.: Feasibility and acceptability of clinic-based telepsychiatry for low-income Hispanic primary care patients. *Telemed. J. e-Health* **18**(4), 297–304 (2012)
30. Ye, J., Shim, R., Lukaszewski, T., et al.: Telepsychiatry services for Korean immigrants. *Telemed. J. e-Health* **18**(10), 797–802 (2012)
31. Shore, J.H., Brooks, E., Savin, D., Orton, H., Grigsby, J., Manson, S.M.: Acceptability of telepsychiatry in American Indians. *Telemed. J. e-Health* **14**(5), 461–466 (2008)
32. Weiner, M.F., Rossetti, H.C., Harrah, K.: Videoconference diagnosis and management of Choctaw Indian dementia patients. *Alzheimer's Dementia* **7**(6), 562–566 (2011)
33. [www.denlilleprins.org](http://www.denlilleprins.org)
34. Mucic, D.: Transcultural telepsychiatry and its impact on patient satisfaction. *J. Tel. Telecare* **16**(5), 237–242 (2010)
35. Parkman, S., et al.: Ethnic differences in satisfaction with mental health services among representative people with psychosis in south London: PRISM study 4. *Br. J. Psychiatry* **171**, 260–264 (1997)
36. Burnett, R., et al.: The first contact of patients with schizophrenia with psychiatric services: social factors and pathways to care in a multi-ethnic population. *Psychol. Med.* **29**(2), 475–483 (1999)
37. Tolmac, J., Hodes, M.: Ethnic variation among adolescent psychiatric in-patients with psychotic disorders. *Br. J. Psychiatry* **184**, 428–431 (2007)
38. Bower, P., Gilbody, S.: Managing common mental health disorders in primary care: conceptual models and evidence base. *Br. Med. J.* **330**, 839–842 (2005)
39. Smith, S.M., Allwright, S., O'Dowd, T.: Effectiveness of shared care across the interface between primary and specialty care in chronic disease management. *Cochrane Database of Systematic Reviews*, p. CD004910 (2007)

40. Kates, N., Ackerman, S.: Shared mental health care in canada. a compendium of current projects (1997)
41. Gask, L., Sibbald, B., Creed, F.: Evaluating models of working at the interface between mental health services and primary care. *Br. J. Psychiatry* **170**, 6–11 (1997)
42. Kelly, B.J. et al.: Shared care in mental illness: a rapid review to inform implementation. *Int. J. Ment. Health Syst.* **5**, 31 (2011). <http://doi.org/10.1186/1752-4458-5-31>
43. Psykiatrifonden (2012). <http://www.psykiatrifonden.dk/forside/nyheder/seneste-nyheder-2/om-undersogelsen-af-lediges-psykiske-problemer>
44. Udviklingen i udlndinges integration i det danske samfund. Rapport. (The development of foreigners' integration into the Danish society. 2006. Report. In Danish). Copenhagen, Denmark: The Ministry of Refugees, Immigrants and Integration of Denmark (2006)
45. Mucic, D.: International telepsychiatry: a study of patient acceptability. *J. Telemed. Telecare* **14**, 241–243 (2008)
46. Hilty, D.M., Yellowlees, P.M., Tarui, N., et al.: Mental Health Services for California American Indians: Usual Service Options and a Description of Telepsychiatric Consultation to Select Sites. In: Ramesh, Madhavan, Shahram, Khalid (eds.) *Telemedicine*, pp. 75–104. Press, InTech Open (2013)
47. Yellowlees, P.M., Hilty, D.M., Mucic, D.: Global/Worldwide e-Mental Health: International and Futuristic Perspectives of Telepsychiatry and the Future. In: Mucic, D., Hilty, D.M. (eds.) *e-Mental Health*, pp. 233–249. Springer, Cham (2016)

**Part II**

**Biomedical Engineering, Trends, Research  
and Technologies**

# A Comparison of Performance of Sleep Spindle Classification Methods Using Wavelets

Elena Hernandez-Pereira, Isaac Fernandez-Varela  
and Vicente Moret-Bonillo

**Abstract** Sleep spindles are transient waveforms and one of the key features that contributes to sleep stages assessment. Due to the large number of sleep spindles appearing on an overnight sleep, automating the detection of this waveforms is desirable. This paper presents a comparative study over the sleep spindle classification task involving the discrete wavelet decomposition of the EEG signal, and seven different classification algorithms. The main goal was to find a classifier that achieves the best performance. The results reported that Random Forest stands out over the rest of models, achieving an accuracy value of  $94.08 \pm 2.8$  and  $94.08 \pm 2.4\%$  with the symlet and biorthogonal wavelet families.

**Keywords** Sleep spindles · Wavelets · Machine learning

## 1 Introduction

According to the current AASM definition [3], the Sleep Spindle (SS) is a “train of distinct waves with frequency 11–16 Hz (most commonly 12–14 Hz) with a duration greater or equal to 0.5 s, usually maximal in amplitude in the central derivations”. The sleep spindle waves are characterized by progressively increasing then gradually decreasing amplitude, that may be present in low voltage background Electroencephalogram (EEG), superimposed to delta activity, or temporally locked to a vertex sharp wave and to a K complex [16]. Spindles are one of the key features that contributes to sleep stages assessment, specifically is one of the hallmarks of Non-Rapid

---

E. Hernandez-Pereira (✉) · I. Fernandez-Varela · V. Moret-Bonillo  
Faculty of Informatics, Department of Computer Science,  
University of A Coruña, Campus de Elviña s/n, 15071 A Coruña, Spain  
e-mail: elena.hernandez@udc.es

I. Fernandez-Varela  
e-mail: isaac.fvarela@udc.es

V. Moret-Bonillo  
e-mail: vicente.moret@udc.es

Eye Movement (NREM) stage 2 sleep, both in adults and children. Unfortunately, their visual identification is very time-consuming (there are typically hundreds of sleep spindles in a full night recording), and they are borderline in frequency or duration, or superimposed on other waveforms. Moreover, there are varying definitions of sleep spindle in the literature, making the criterion used for spindle scoring inconsistent across studies. Another limitation is that interscorer reliability for visual identification suggests a variability between scorers possibly due to subjectivity or expertise of the scorer [27]. Thus, automated sleep spindle detectors have been developed to reduce the workload of experts and eliminate the subjectivity.

The earliest sleep spindle detectors were dependent upon hardware [13, 20]. After these detection systems several software solutions have been attempted. Two principal approaches become accepted: those using band-pass filtering and amplitude detection, and those applying feature extraction follow by decision-making for classification. The first approach, followed in [10, 29], suffers from the interscorer variability and that is one of the reasons for the second approach to be a noted research line. Concerning algorithms based on features extraction followed by classification, the Short Time Fourier Transform (STFT) is a suitable tool to identify the frequency content of the sleep spindles. In [17], Gorur used STFT coefficients as inputs of a classifier. An agreement rate of 88.7 and 95.4 % were obtained with a multilayer perceptron (MLP) and a support vector machine (SVM) respectively. Another method used for features extraction is adaptive autoregressive modelling (AAR). In [1] the AAR coefficients were used as inputs for different classifiers: a discrete perceptron, a MLP and a SVM. The results obtained were compared in terms of sensitivity, achieving values of 99.2 %, 89.1 % and 94.6 % respectively. In recent years advanced time-frequency analysis tools like wavelets have been applied to the sleep EEG to derive improved feature vectors for sleep spindles. Ahmed et al. [2] proposed a automatic detector based upon the Teager Energy Operator (TEO) and Wavelet Packet Energy Ratio, and achieved an accuracy of 93.9 %. In [11] a multi-resolution decomposition technique based on wavelets and STFT, is developed to detect sleep spindles. After the detection, TEO is applied to determine spindle duration. By this approach, an overall sensitivity and specificity of 96.17 and 95.54 % were achieved. TEO is employed too in [19] where this operator isolated candidate spindle zones on sleep EEG and spectral edge frequency confirmed its presence. The algorithm used a normalized threshold and did not require patient-specific adjustments. It achieved 80 % and 97.6 % values of sensitivity and specificity respectively. Günes et al. [18] proposed a hybrid method based on time and frequency domain features. Welch spectral analysis has been used for the extraction of frequency domain features and a MLP for classification. The obtained classification accuracies for three feature sets (only time domain, only frequency domain and both frequency and time domain features) were 100, 56.86 and 93.84 %. In [24], an algorithm that models the amplitude frequency spindle distribution with a bivariate normal distribution is proposed. Spindle detection is not directly based on amplitude and frequency thresholds, but instead on a spindle distribution model that is automatically adapted to each individual subject. Authors concluded that normal modelling enhanced performance and improved spindle detection quality.

This work studies the capabilities of several machine learning techniques to classify sleep spindles. The feature extraction is accomplished using a discrete wavelet decomposition applied to the raw samples of the EEG signal segments. The paper is structured as follows: Sect. 2 proposes the research methodology, Sect. 3 describes the experimental procedure used in the research, Sect. 4 presents the results obtained and finally, the conclusions are presented in Sect. 5.

## 2 Research Methodology

The main objective of this work is to obtain a method that achieves the best accuracy results in the sleep spindle classification task. Over the EEG signal from several sleep recordings, a set of isolated waveforms was obtained. Using these patterns, the coefficients of a discrete wavelet decomposition were used as inputs for several classifiers.

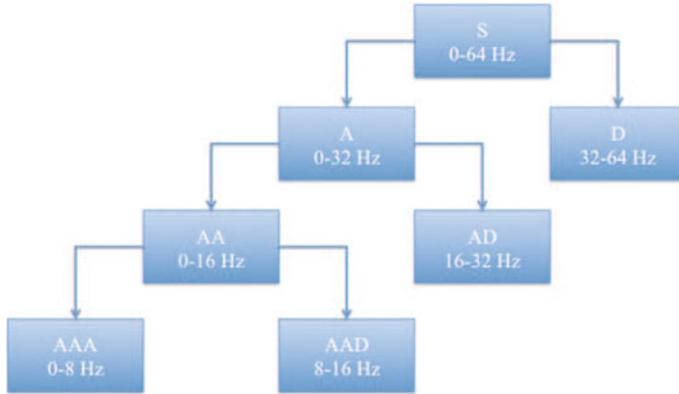
### 2.1 Data Set

Patient data was gathered from the Sleep Laboratory of the André Vésale Hospital in Belgium. It consists of eight whole-nights recordings coming from patients—4 men and 4 women aged between 31 and 53—with different pathologies. Two EOG channels, three EEG channels and one submental EMG channel were recorded. The sampling frequency was 200 Hz for six records of the complete data set, 100 and 50 Hz for the two remaining ones. A segment of 30 min was extracted from each night from the central EEG channel for spindles scoring. No effort was made to select good spindle epochs or noise free epochs, in order to reflect reality as well as possible. These segments were given to a medical expert for sleep spindle scoring. The total number of identified spindles was 289 [10].

### 2.2 Feature Extraction

The wavelet transform is an efficient tool for decomposing a signal into a fundamental function set and obtaining sub-band localization. Figure 1 depicts the wavelet decomposition tree.

In the first step, a high pass filter  $g(n)$  and a low pass filter  $h(n)$  are applied to the original signal  $x(n)$ . After the filtering process, half of the samples at high frequency are discarded according to Nyquist Criteria. This operation is performed recursively for every remaining sample and the desired frequency intervals are obtained. We can mathematically express this procedure as follows:



**Fig. 1** Wavelet decomposition tree

$$Y_{high}[k] = \sum x[n]\dot{g}[2k - n] \quad (1)$$

$$Y_{low}[k] = \sum x[n]\dot{h}[2k - n] \quad (2)$$

where  $Y_{high}[k]$  and  $Y_{low}[k]$  are the outputs of the high pass (D) and low pass (A) filters, respectively.

The discrete wavelet transformation [26] provides a decomposition of a given signal into a set of approximation ( $a_i$ ) and detail ( $d_i$ ) coefficients of level  $i$ . The decomposition process can be iterated, with successive approximations being decomposed in turn, so that a signal is broken down into many lower-resolution components. Thus, in this case the samples of the EEG signal, are processed to obtain a level-1 transformation ( $a_1$  and  $d_1$  coefficients). Subsequently, each set of  $a_i$  coefficients is decomposed into a set of approximation  $a_{i+1}$  and detail  $d_{i+1}$  coefficients. Also, to obtain this decomposition some different types of wavelets functions can be used. The level-detail was determined taking into account the sample rate of the EEG signal and the wavelet families were chosen after performing some other experiments and discarding several wavelet families, specifically, the Symlet, Haar, Daubechies, Coiflets, Biorthogonal, and the discrete approach of the Meyer wavelet [9].

## 2.3 Classification

In this section, we provide an overview of the methods used in the research for sleep spindle classification. Several approaches were considered, two lineal models—a one-layer feedforward neural network and a proximal support vector machine—, and five non linear ones—a multilayer feedforward neural network, a classification tree, a Random Forest, a Support Vector Machine and a Naive Bayes classifier—.

- One-layer Feedforward Neural Network, One-lay FNN

The one-layer feedforward neural network (FNN) is a single-layer FNN without hidden layers. This is a linear classification system that was trained using the supervised learning method proposed in [8]. The contribution of this learning method is that it is based on the use of an alternative cost function that measures the errors *before* the nonlinear activation functions instead of *after* them, as is normally the case. An important consequence of this formulation is that the solution can be obtained directly using a system of linear equations due to the fact that the new cost function is convex [14]. So, the method avoids local minima, and a very good approximation to the global minimum of the error function is obtained.

- Multilayer Feedforward Neural Network, FNN

The multilayer feedforward neural network is one of the most commonly used neural network classification algorithms [4]. The architecture used for the classifier consisted of a three layer feed-forward neural network: two hidden and one output layer. The optimal number of hidden neurons for this problem was empirically obtained.

- Classification Trees, Class. Tree

Classification trees are used to predict membership of cases or objects in the classes of a categorical dependent variable from their measurements on one or more predictor variables. In these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels [5]. Each internal (non-leaf) node of the tree is labelled with an input feature. The arcs coming from a node labelled with a feature are labelled with each of the possible values of the feature. Each leaf of the tree is labelled with a class or a probability distribution over the classes. A tree can be “learned” by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner. The recursion is completed when the subset at a node has all the same value of the target variable, or when splitting no longer adds value to the predictions. This process of top-down induction of decision trees is by far the most common strategy for learning decision trees from data [25].

- Random Forests, RF

Random Forests [7] are an ensemble learning method for classification that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes. For an ensemble of decision trees for a multiclass classification function, one of the general methods is Bagging. This method is the simpler, more robust and more highly parallel technique. In the Bagging version used, a fixed-sized fraction of the training data is employed to construct each classifier in the ensemble. The Bagging method simply produces an ensemble of N decision trees constructed from N random subsets of the training data, where each subset is of the fixed-size mentioned in the previous sentence. With Bagging, the original method from the literature [6] of choosing a subset of points from a complete training set of N points was to choose a *bootstrap* sample [12]. Simply put, this means randomly choosing N points with equal probability from the set with replacement, so that some points may be chosen more than once or not at all.

To compute prediction of an ensemble of trees for unseen data, the Random Forest model takes an average of predictions from individual trees. To estimate the prediction error of the bagged ensemble, predictions for each tree are computed on its out-of-bag observations, are averaged over the entire ensemble for each observation and then the predicted out-of-bag response is compared with the true value at this observation.

- **Support Vector Machine, SVM**

A Support Vector Machine is a supervised classification technique that works by nonlinearly projecting the training data in the input space to a feature space of higher (infinite) dimension by the use of a kernel function. This results in a linearly separable data set by a linear classifier. In many instances, classification in high dimension feature spaces results in overfitting in the input space; however, in SVMs, overfitting is controlled through the principle of structural risk minimization [28]. The empirical risk of misclassification is minimized by maximizing the margin between the data points and the decision boundary [21].

- **Naive Bayes, NB**

Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features. This assumption dramatically reduces the number of parameters that must be estimated to learn the classifier. Naive Bayes is a widely used learning algorithm, for both discrete and continuous inputs. The Naive Bayes Classifier technique is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods [23].

- **Proximal Support Vector Machine, pSVM**

The proximal Support Vector Machine [15] is a method that classifies points assigning them to the closest of two parallel planes (in input or feature space) that are pushed as far apart as possible. The difference with a SVM is that this one classifies points by assigning them to one of two disjoint half-spaces. The pSVM leads to an extremely fast and simple algorithm by generating a linear or nonlinear classifier that merely requires the solution of a single system of linear equations.

### **3 Experimental Procedure**

In order to characterize the performance of the system, the sensitivity and accuracy standard measures were used. The procedure presented has two stages: a feature extraction stage that establishes the main inputs for classification and the classification model itself. The initial step of the proposed methodology is the processing of the available EEG signals, obtaining isolated waveforms. The positive examples were identified by the medical expert and in order to achieve a balanced data set, the same number of negative examples were selected from the whole set of recordings. To get these negative examples, signal windows of 0.5 s (minimum sleep spindle duration) were randomly selected.

The sleep spindles contain low frequency components that are placed at 11–16 Hz band. Therefore, 8–16 Hz band covers all the information required for spindle detection. Thus, the computational cost and detection errors can be reduced by limiting the search to this sub-band. A discrete wavelet transformation was used as a pre-processing phase to reduce and fix the number of inputs of the classifier. Experimentally, for this work it was determined that the absolute value of the level-4-detail coefficients ( $d_4$ ), the level-3-detail ( $d_3$ ), and the level-2-detail ( $d_2$ ) are the set of inputs that obtains the best sleep spindle classification results for the different sample rates (200 Hz, 100 Hz and 50 Hz respectively). Also, in order to make the decomposition, the symlet and biorthogonal wavelet families were used with a length of the filter equal to 10, 14 and 7 (symlet of order:  $O_5$ , sym5 and  $O_7$ , sym7; and biorthogonal of order  $O_{1,5}$ , bio1.5).

The number of coefficients supplied by the wavelet transformation depends on the number of samples of the supplied pattern. In this case, the minimum number of samples is 25 (as 0.5 s is the minimum duration of a sleep spindle and the lowest sample rate is 50 Hz), therefore, the number of corresponding wavelet coefficients is 13. As the number of inputs to the classification module must be fixed, the maximum number of coefficients that could be used is 13. For those patterns in which the duration is the minimum (0.5 s) all the coefficients (13) were used. For those other patterns with a duration greater than 0.5 s (for which more than 13 coefficients could be obtained) the first 13 coefficients were used as inputs to the classifier.

The experimental procedure is detailed as follows:

1. Extract the initial set of features to be used as inputs.
2. For each nonlinear classifier, establish its architecture. For the FNN a two hidden layer architecture with 10 and 8 units respectively was chosen. For the Random Forest, the number of trees chosen was 20 and for the SVM, the RBF kernel function was used.
3. Take the whole data set and generate 10-fold cross validation sets in order to better estimate the true error rate of each model. Eight folds are used to train the models, and the remainder ones to validate and test them respectively.
4. Train each model and obtain 10 performance measures over the validation sets and the test sets.
5. Select the best model in terms of accuracy.

The experiments performed in this work were executed using the software tool Matlab [22].

## 4 Results

In this section, the results obtained over the test set, after applying wavelet transformation and several classifiers are shown and compared in terms of accuracy and sensitivity. These results are yield against the standard reference, i.e. the medical expert scores. Tables 1 and 2 show the performance measures obtained.

**Table 1** Sleep spindle classification results

	pSVM	One-lay. FNN	Class. Tree	RF	FNN	SVM	NB
sym5	$86.20 \pm 3.9$	$89.58 \pm 3.9$	$91.83 \pm 4.5$	<b><math>94.08 \pm 2.8</math></b>	$87.89 \pm 4.5$	$89.01 \pm 4.3$	$93.66 \pm 1.7$
sym7	$83.66 \pm 6.8$	$85.35 \pm 4.6$	$86.20 \pm 3.2$	<b><math>89.58 \pm 3.2</math></b>	$82.54 \pm 6.6$	$86.90 \pm 4.5$	$85.77 \pm 5.2$
bio1.5	$86.20 \pm 2.8$	$88.59 \pm 2.7$	$93.38 \pm 4.2$	<b><math>94.08 \pm 2.4</math></b>	$86.48 \pm 6.4$	$88.59 \pm 3.9$	$93.66 \pm 2.9$

Mean test set accuracy (%) of a 10-fold cv. Best values marked in bold font

**Table 2** Sleep spindle classification results

	pSVM	One-lay. FNN	Class. Tree	RF	FNN	SVM	NB
sym5	$97.99 \pm 1.9$	$96.01 \pm 3.4$	$91.86 \pm 6.8$	$95.78 \pm 2.8$	$86.45 \pm 6.1$	$92.67 \pm 4.3$	$95.15 \pm 1.5$
sym7	$95.04 \pm 4.9$	$90.30 \pm 3.8$	$85.47 \pm 7.8$	$88.69 \pm 6.5$	$82.16 \pm 7.9$	$89.67 \pm 4.5$	$89.63 \pm 4.3$
bio1.5	$99.72 \pm 0.8$	$98.89 \pm 1.1$	$93.14 \pm 4.3$	$95.81 \pm 3.3$	$87.05 \pm 7.9$	$92.08 \pm 4.1$	$96.36 \pm 3.6$

Mean test set sensitivity (%) of a 10-fold cv

Among the linear models tested (pSVM and one-layer FNN), the one-layer FNN showed the best performance, achieving the highest accuracy for all the wavelet families used. For this classifier, the bio1.5 wavelet offers the best inputs. Over the non-linear models, the Random Forest obtained the best results. These facts state no matter what wavelet family used. Nevertheless, the biorthogonal wavelet is the one that provides the best inputs for the classifier.

In terms of sensitivity, the linear models, pSVM and one-layer FNN, showed the highest values with the bio1.5 wavelet, but their accuracy values are not as good as expected. For the Random Forest model, the sensitivity values achieved were satisfactory for the sym5 and bio1.5 wavelets.

## 5 Conclusions

This paper presents a comparative study over the sleep spindle classification task involving the discrete wavelet decomposition of the EEG signal, and seven different classification algorithms. The main goal was to find a classifier that achieves the best accuracy results.

As a starting point, the extraction of isolated waveforms was carried out. Up to the authors knowledge, not many previous methods were proposed for sleep spindle classification that used the discrete wavelet decomposition as the feature extraction method. In this environment, several wavelets families were probed, being the symlet and biorthogonal families the ones that obtain the best results for the classifiers.

The results obtained were similar to those reported in the bibliography [2, 11] but a fair comparative study is not possible due to differences in both datasets and evaluation methods. In this work, from the classifier point of view, the results reported

that Random Forest is the best option, achieving an accuracy value of  $94.08 \pm 2.8$  and  $94.08 \pm 2.4\%$  with the symlet (order  $O = 5$ ) and biorthogonal (order  $O = 1.5$ ) wavelet families. For these models, the sensitivity values are similar ( $95.78 \pm 2.8$  and  $95.81 \pm 3.3$  respectively). The results are encouraging and a deeper study will be done first in the negative examples extraction task. Different waveforms durations should be considered to make more difficult the classifier task instead of providing it with easy examples. Besides, we plan to test the use of ensembles of classifiers, trying to take advantage of the strengths of the different algorithms tested here and combine them in order to improve the classification accuracy. Finally, to confirm Random Forest best results, experiments over the entire signal length should be performed.

**Acknowledgments** This research was partially funded by the Xunta de Galicia (Grant code GRC 2014/035) and by the Spanish Ministerio de Economía y Competitividad, MINECO, under research project TIN2013-40686P both partially supported by the European Union ERDF.

## References

1. Acir, N., Güzelis, C.: Automatic recognition of sleep spindles in EEG by using artificial neural networks. *Expert Syst. Appl.* **27**(3), 451–458 (2004)
2. Ahmed, B., Redissi, A., Tafreshi, R.: An automatic sleep spindle detector based on wavelets and the teager energy operator. In: Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2596–2599 (2009)
3. Berry, R.B., et al.: The AASM Manual for Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. American Academy of Sleep Medicine, Darien, Illinois (2015)
4. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, New York (1995)
5. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Chapman & Hall, New York (1984)
6. Breiman, L.: Bagging predictors. *Mach. Learn.* **24**(2), 123–140 (1996)
7. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
8. Castillo, E., Fontenla-Romero, O., Alonso-Betanzos, A., Guijarro-Berdiñas, B.: A global optimum approach for one-layer neural networks. *Neural Comput.* **14**(6), 1429–1449 (2002)
9. Daubechies, I.: Ten lectures on wavelets. In: Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (1992)
10. Devuyst, S., Dutoit, T., Stenuit, P., Kerckhofs, M.: Automatic sleep spindles detection. Overview and development of a standard proposal assessment method. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC (2011)
11. Duman, F., Erdamar, A., Erogul, O., Telatar, Z., Yetkin, S.: Efficient sleep spindle detection algorithm with decision tree. *Expert Syst. Appl.* **36**(6), 9980–9985 (2009)
12. Efron, B.: Bootstrap methods: another look at the jackknife. *Ann. Stat.* **7**, 1–26 (1979)
13. Fish, D., Allen, P., Blackie, J.: A new method for the quantitative analysis of sleep spindles during continuous overnight eeg recordings. *Electroencephalogr. Clin. Neurophysiol.* **70**(3), 273–277 (1988)
14. Fontenla-Romero, O., Guijarro-Berdiñas, B., Pérez-Sánchez, B., Alonso-Betanzos, B.: A new convex objective function for the supervised learning of single-layer neural networks. *Pattern Recognit.* **43**(5), 1984–1992 (2010)

15. Fung, G., Mangasarian, O.: Proximal support vector machine classifiers. In: Provost, F., Srikanth, R. et al. (eds.) *Proceedings KDD-2001: Knowledge Discovery and Data Mining*. pp. 77–86. San Francisco, CA, Association for Computing Machinery, New York (2001)
16. Gennaro, L.D., Ferrara, M.: Sleep spindles: an overview. *Sleep Med. Rev.* **7**(5), 423–440 (2003)
17. Görür, D.: Automated Detection of Sleep Spindles. MSc thesis (2003)
18. Güneş, S., Dursun, M., Polat, K., Yosunkaya, C.: Sleep spindles recognition system based on time and frequency domain features. *Expert Syst. Appl.* **38**(3), 2455–2461 (2011)
19. Imtiaz, S.A., Saremi-Yarahmadi, S., Rodriguez-Villegas, E.: Automatic detection of sleep spindles using teager energy and spectral edge frequency. In: *Biomedical Circuits and Systems Conference (Biocas)*, 2013 IEEE, pp. 262–265 (2013)
20. Kumar, A., Hofman, W., Campbell, K.: An automatic spindle analysis and detection system based on the evaluation of human ratings of the spindle quality. *Waking Sleep*. 325–333 (1979)
21. Mashao, D.: Comparing SVM and GMM on parametric feature-sets. In: *Proceedings of the 15th Annual Symposium of the Pattern Recognition Association of South Africa* (2004)
22. MATLAB: version 8.4.0.150421 (R2014b). The MathWorks Inc., Natick, Massachusetts (2014)
23. Mitchell, T.: *Machine Learning*. McGraw Hill (1997)
24. Nonclercq, A., Urbain, C., Verheulpen, D., Decaestecker, C., Bogaert, P.V., Peigneux, P.: Sleep spindle detection through amplitude? Frequency normal modelling. *J. Neurosci. Methods* **214**(2), 192–203 (2013)
25. Quinlan, J.R.: Induction of decision trees. *Mach. Learn.* **1**, 81–106 (1986)
26. Rao, R.M., Bopardikar, A.S.: *Wavelet Transformations. Introduction to Theory and Applications* (1998)
27. Ray, L.B., Fogel, S.M., Smith, C.T., Peters, K.R.: Validating an automated sleep spindle detection algorithm using an individualized approach. *J. Sleep Res.* **19**(2), 374–378 (2010)
28. Vapnik, V.: *Statistical learning theory. Adaptive and learning systems for signal processing, communications, and control* (1998)
29. Ventouras, E.M., Monoyiou, E.A., Ktonas, P.Y., Paparrigopoulos, T., Dikeos, D.G., Uzunoglu, N.K., Soldatos, C.R.: Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: a feasibility study. *Comput. Methods Progr. Biomed.* **78**(3), 191–207 (2005)

# Evaluation of Head-Mounted Displays for Macular Degeneration: A Pilot Study

Howard Moshtael, Lanxing Fu, Ian Underwood and Baljean Dhillon

**Abstract** Head-mounted displays are increasingly available and have been proposed as a platform for a new class of low vision aid. Two kinds of head-mounted display, smart glasses and a smartphone-based headset, are evaluated here in terms of their visibility and desirability for users with age-related macular degeneration. This evaluation is performed through a test to measure the extent of visibility of the screens, through a reading speed test and through a questionnaire. All but one of the participants could read from the displays and see at least 45 % of at least one of the screens. The majority found it easier to read from the displays than from paper.

**Keywords** Head-mounted displays • Low vision • Visual aids • Image processing • Reading

## 1 Introduction

The head-mounted display (HMD) is a niche but growing class of electronic display system. An HMD comprises a miniature electronic display screen in close proximity to one or both eyes viewed through compact imaging optics which cause a highly magnified virtual image of the miniature screen to appear at a comfortable

---

H. Moshtael (✉)

School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK  
e-mail: hfm30@hw.ac.uk

L. Fu

Ophthalmology Department, Royal Blackburn Hospital, Blackburn, UK  
e-mail: lfu@doctors.org.uk

I. Underwood

School of Engineering, University of Edinburgh, Edinburgh, UK  
e-mail: i.underwood@ed.ac.uk

B. Dhillon

School of Clinical Sciences, University of Edinburgh, Edinburgh, UK  
e-mail: baljean.dhillon@ed.ac.uk

distance for the viewer. A simple HMD can be made by mounting a smartphone into a frame which incorporates the necessary optics at a suitable distance. This idea was popularized by Google Cardboard and can be used as a budget “virtual reality” system. Another class of HMD features a partially transparent screen instead of the typical opaque type. As such, the user’s view of their surroundings is not occluded. When the display screen is housed in something resembling a spectacle frame it is referred to as “smart glasses”. This technology continues to be actively developed with many products currently available only as developer kits. But with numerous vendors set to release commercial products the technology is expected to become much more widespread; the smart glasses market is expected to almost double in the next 4 years [1].

Age-related macular degeneration (AMD) is one of the leading causes of blindness in the Western world [2]. It is characterized primarily by a scotoma in the central visual field and thus can severely inhibit daily tasks [3]. Continuing to be able to read is a central concern to those that develop AMD [4]. Various low vision aids (LVAs) exist to enhance the use of residual vision, with the most common technique being magnification.

The use of HMDs in LVAs was first attempted while the technology was in an early stage of its development, meaning the resulting device was bulky and 1 kg in weight [5, 6]. Later a few devices were commercialized and a 2004 study compared four of them with conventional optical LVAs [7]. They were found to enhance distance and intermediate visual acuity, but to offer no significant improvement for near acuity. Advances in display technology have improved on the known limitations of these devices, such as their restricted near field of view and substantial bulk [8]. Some recent work on utilizing HMDs in LVAs includes: A well-known augmented reality device, Google Glass, as a platform for edge enhancement [9]; a depth camera connected to an HMD as a navigation aid [10]; non-linear magnification to enhance reading [11]. In parallel to the development of hardware, various image processing methods for the specific benefit of the visually impaired have been developed and begun to be tested [12].

To evaluate the suitability of an HMD for low vision users, we need to measure the extent to which they can see and read from the display, as well as their subjective feedback using it. This data is lacking in the literature. Further, the methodology for making these measurements on displays in general has not been developed. This paper investigates a device from both of the HMD classes described above: Partially transparent smart glasses and an opaque smartphone-based system. It presents the results of a pilot trial to measure the area of the screen visible to the low vision users, their ability to read from the screens and their experience using the devices.

## 2 Methodology

### 2.1 Participants

Ten individuals with AMD were recruited from the low vision clinic of the Princess Alexandra Eye Pavilion, Edinburgh, UK, and from the Macular Society. Relevant details are included in Table 1. The age range was between 52 and 91 years with equal numbers of both sexes. All patients were English speaking. The planned study received a favorable opinion from the National Research Ethics Service Committee and gained R&D approval from NHS Lothian. Informed, written consent was obtained from all participants prior to commencement of the study.

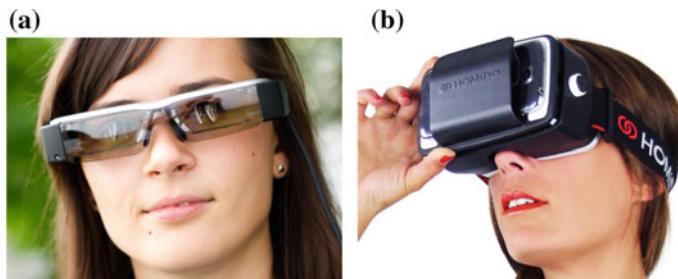
### 2.2 Head-Mounted Displays

The model of smart glasses used in the study was the Epson Moverio BT-200 (Seiko Epson, Japan; see Fig. 1a). It is comprised of two miniature high definition color displays, one aligned to each eye, which then appear as a single screen in the center of the visual field with a field of view of approximately 23°. The screen appears at infinity. For the other HMD, a popular smartphone with a high pixel density of 538 ppi—the LG G3 (LG Electronics Inc., South Korea)—was chosen. The Homido headset (Homido, France), shown in Fig. 1b, holds the smartphone and uses a pair of lenses to magnify and image the screen, and helpfully includes a means to adjust the interpupillary distance. It includes three pairs of lens mounts that are chosen for normal, myopic or hyperopic vision. The field of view is approximately 100° and the image appears close to the eye.

**Table 1** Clinical characteristics and past experience with LVAs

Participant	Sex	Age	Years since diagnosis	Sight registration status	Wet/dry AMD	Reading acuity (logRAD score <sup>a</sup> )	Previous use of optical LVAs?	Previous use of electronic LVAs?
101	M	59	15	Blind	Dry	1.3	Y	Y
102	F	89	5	Unregistered	Dry	0.435	Y	N
103	M	52	5	Unregistered	Not told	0.73	Y	N
104	M	78	20	Unregistered	Dry	0.8	Y	N
105	F	80	14	Unregistered	Both	0.5	Y	N
106	M	74	20	Blind	Wet	1.56	Y	Y
107	F	80	11	Blind	Wet	1.67	Y	Y
108	M	76	8	Unregistered	Dry	0.725	Y	N
109	F	74	13	Blind	Both	>1.7	Y	Y
110	F	91	25	Unregistered	Wet	0.525	Y	N

<sup>a</sup>logRAD is the reading equivalent of logMAR; see Sect. 2.4



**Fig. 1** Devices used **a** Epson Moverio BT-200 with the lighter of the two neutral density filters [idnes.cz]. **b** Homido headset with smartphone [homido.com]

The luminance was measured through the lens at the position where the eye is placed using a Topcon BM-9 luminance meter (Topcon Technohouse Corporation, Japan). At the phone's 80 % brightness setting, as used in the study, the luminance was measured to be  $190 \pm 3 \text{ cd/m}^2$  for an all-white screen and less than  $1 \text{ cd/m}^2$  for an all-black screen. The luminance of a white and black screen on the smart glasses, set to maximum brightness, was measured to be  $400 \pm 50 \text{ cd/m}^2$  and  $6 \pm 1 \text{ cd/m}^2$  respectively. The larger standard deviation arises due to the relative difficulty of measuring the brightness for the very small microdisplay screen. As the smart glasses are partially transparent it was also necessary to maintain a consistent background illumination level. A black background was chosen with a reflected luminance of  $2 \pm 1 \text{ cd/m}^2$ . The luminance reflected from the white page from which the print was read was dependent on the angle at which it was tilted due to the light source being a ceiling bulb. It was measured as  $70 \pm 3 \text{ cd/m}^2$  when vertical but  $92 \pm 3 \text{ cd/m}^2$  when tilted  $20^\circ$  from the vertical towards the ceiling.

Each participant used both HMDs, with the order of use randomized. When using the smart glasses, participants wore their distance glasses, whereas in the case of the smartphone headset the optimal lens mount was selected by each participant.

### 2.3 Extent of Visibility

As AMD involves the appearance and growth of scotomas, a test was needed to measure which parts of the screen were and were not visible to the participants. We refer to the percentage of the screen seen by patients as the extent of visibility. We used the approach taken in perimetry tests, the standard method for measuring the visual field, but adapted for the HMD. The test involves a flashing cross at the center of the screen to be the point of fixation for the participant. The stimuli, white dots of approximately 5 arcminutes in diameter, were flashed for a duration of 300 ms and the subject was asked to say "Yes" when they observed one. The time window for responding was designed to automatically adapt to the speed with which the participant would respond. A second chance was given to all stimuli that

were not seen the first time. For the smart glasses, a total of 29 stimuli were evenly spread across the screen, whereas for the smartphone the 24 stimuli were evenly spread across the central 60° field of view. The duration of the test is between 1 and 4 min.

Though similar in method to perimetry, this test does not seek to measure the visual field of the participants. Rather, it maps the areas of the screen visible and not visible to the participant. This data could be used to tailor the presentation of screen content to the particular vision of the user. Further research is required to investigate this.

## 2.4 *Reading Performance*

Performance reading from the smart glasses and smartphone respectively were compared to performance reading from print. The English Radner Reading Chart, held at 25 cm, was used to measure reading acuity and reading speed from print. Those who could not read the top line at 25 cm were allowed to read from a distance of their choosing; this distance was then measured and the text size calculated accordingly. It consists of a bank of 28 sentences, each with a standardized structure.

The size of the print is defined by the height of a lower case “x” character [13]. The minimum angle of resolution (MAR) is defined as the angle in arcminutes subtended by a fifth of this height (the stroke width of a letter). The print sizes are logarithmically scaled and range from 1.3 to 0 logMAR when viewed at a distance of 25 cm. The Radner Chart defines logRAD (reading acuity determination) score for measuring reading acuity as logRAD (equivalent to logMAR) of the smallest sentence partially read plus 0.005 for each syllable incorrectly read.

As the HMDs use a virtual display, the print size is measured in terms of visual angle then converted to logMAR. The smartphone based HMD is a simple magnifying glass optical system with the lens held at the focal length from the display. As the eye is positioned with negligible distance from the lens, the visual angle subtended on the display is given by simple trigonometry. In the case of the smart glasses, the optical system is more complicated. However, as the total visual angle subtended from corner to corner of the display is known to be 23° (as stated by the manufacturer), the visual angle of the letters was calculated as a proportion of this.

A limitation to the smartphone-based system is that the smallest text size that can be displayed is 0.5 logMAR. This is because the smartphone display is highly magnified thus the text size on the screen itself has to be very small. Thus even a screen with a very high pixel density display such as the LG G3 used here, the stroke width of the letters at 0.5 logMAR is smaller than a pixel. A limitation of the smart glasses is the relatively small field of view. This means that a sentence with text size 1.1 logMAR or larger would not fit on the screen in the same format presented on the Radner Reading Chart.

## 2.5 Questionnaire

Of equal importance to the objective measures of the participants' performance are their subjective views of using the devices. Therefore a questionnaire was devised to gather this feedback, the questions of which are reproduced in Table 2. The style and format of the questionnaire was based on a validated visual function questionnaire [14]. The questions relate either to the display itself or to the look and feel of the headset. Open questions were also included to gain further feedback. The questionnaire was interviewer-administered. All ten participants took the questionnaire on the smart glasses, and nine took it on the smartphone based system.

## 3 Results

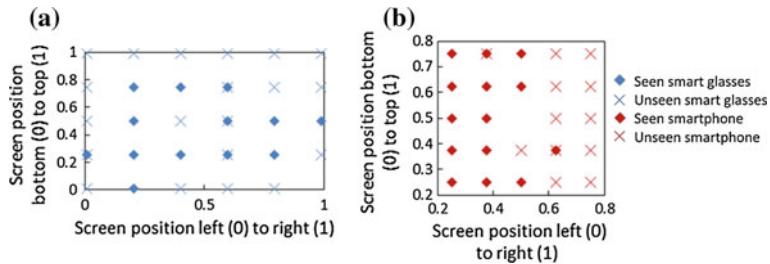
### 3.1 Extent of Visibility

A unique profile was generated for each of the eight participants (103–110) who took the test. The extent of visibility of each screen to each participant was calculated as the percentage of points they saw. At least 45 % of the points were seen on at least one display by all the participants, except participant 109 who saw just one point. All participants saw a higher proportion of points on the smartphone headset than they did on the smart glasses.

By way of example, the results of the test are shown for participant 106 (who is registered blind) using the smart glasses in Fig. 2a and the smartphone headset in Fig. 2b. All the stimuli are shown, be it with a diamond to indicate a point that was seen, or a cross to indicate that it was not. A few stimuli are marked with both, meaning they were not seen on one occasion but were seen on the other. Note that the two plots would not be expected to match because they correspond with different screens of different fields of view (not directly with the retina of the patient).

**Table 2** Questionnaire items and responses

Question	Response options
Compared to reading large print from paper, did you find reading from the display to be...	Much easier/a little easier/the same/a little harder/much harder
Did you prefer reading from the display or from paper?	The display/paper
How comfortable did you find the headset to wear? Was it...	Very comfortable/fairly comfortable/fairly uncomfortable/very uncomfortable
Based on the appearance of the headset, how comfortable would you feel to wear the headset at home? Would you feel...	Very comfortable/fairly comfortable/fairly uncomfortable/very uncomfortable

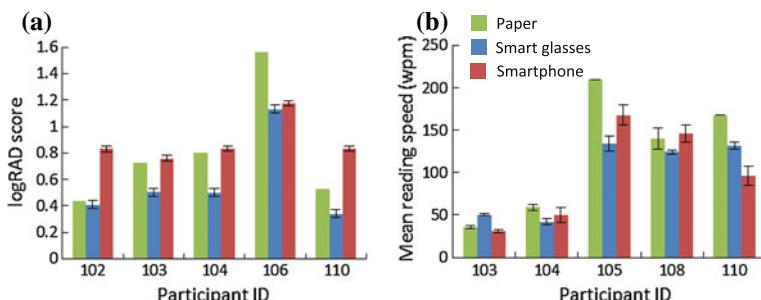


**Fig. 2** Results of the extent of visibility test for participant 106. **a** From the smart glasses. **b** From the smartphone headset

### 3.2 Reading Performance

For a comparison between the three media of paper, smart glasses and smartphone, the participants' vision needed to be good enough to read at least the top line of the chart and bad enough not be able to read below the 0.5 logMAR limit set by the smartphone display. Five participants (102, 103, 104, 106 and 110) fit this criteria and were tested on all three media. Their results on each media are shown in Fig. 3a. The level of uncertainty, arising from the calculation of visual angle, was estimated at  $\pm 0.03$  logRAD for the smart glasses and  $\pm 0.02$  logRAD for the smartphone headset. For all five of these participants, the smart glasses allowed them to read the smallest text size. The mean and standard deviation in logRAD score for paper, smart glasses and smartphone headset are  $0.8 \pm 0.4$ ,  $0.5 \pm 0.3$  and  $0.9 \pm 0.1$ , respectively.

To calculate the mean reading speed for a participant, the speed of reading 5 lines above their critical print size (the print size below which reading speed drops) was averaged together. There were five participants (103, 104, 105, 108 and 110) who read at least 5 lines for each media. Their results for each media are shown in



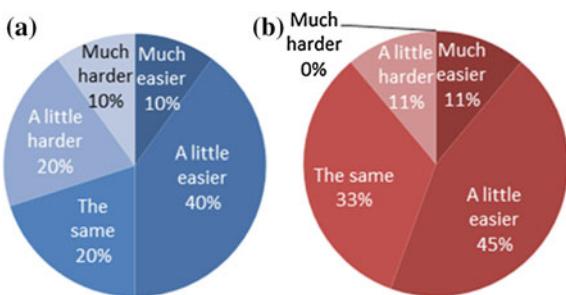
**Fig. 3** Reading performance. **a** The reading acuity for 5 participants, with *error bars* set at  $\pm 0.03$  for smart glasses and  $\pm 0.02$  for smartphone headset. **b** Mean reading speeds for 5 participants, with *error bars* indicating standard error

Fig. 3b. The mean and standard deviation in words per minute on paper, smart glasses and smartphone headset are  $123 \pm 7$ ,  $97 \pm 8$  and  $98 \pm 17$ , respectively.

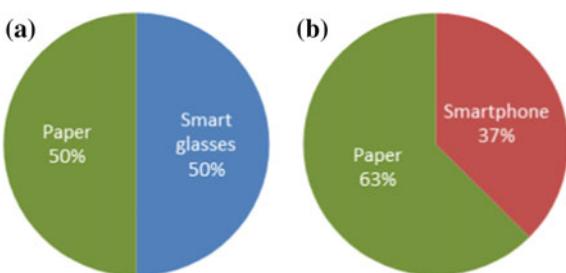
### 3.3 Questionnaire

Responses to the questionnaire (Table 2) are shown in Figs. 4, 5 and 6. Figures 4 and 5 show the responses in terms of the proportion of participants giving a response to each of the options.

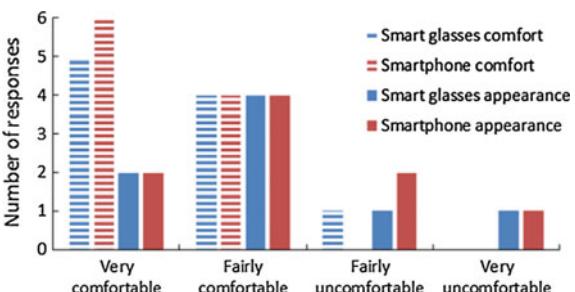
**Fig. 4** Responses to the question, ‘Compared to reading large print from paper, did you find reading from the display to be...’. Results for smart glasses on the *left* and smartphone on the *right*



**Fig. 5** Responses to the question, ‘Did you prefer reading from the display or from paper?’, comparing with **a** smart glasses and **b** smartphone headset



**Fig. 6** Responses to the questions, ‘How comfortable did you find the headset to wear?’ and ‘Based on the appearance of the headset, how comfortable would you feel to wear the headset at home?’



## 4 Discussion

The extent of visibility test, as well as the reading test, sought to determine whether or not AMD would render these screens unusable. It may have been assumed, particularly for the smaller, centrally-positioned screen image of the smart glasses, that a central scotoma would almost entirely block the screen. However, all but two of the participants saw at least 45 % of the points on the smart glasses, and all but one on the smartphone. Also, all but one of the participants were able to read from both displays, including 2 out of 3 who were registered blind. This suggests there is potential for both of these displays to be used for those with AMD.

However, the fact that none of the participants saw every stimulus on the smart glasses suggests that screen tailoring is required for users with AMD. The results in Fig. 2b, for example, show that the area of the screen to the left of the central fixation point was visible to the participant, whereas the area to the right was not. This participant is highly trained in eccentric viewing and thus aligned his preferred retinal locus, located in the top right of his vision, to the fixation point (rather than the fovea as for normal viewing). Consequently his visual field was shifted towards the bottom left, thus partly accounting for the results. This emphasises how the results are relative to the fixation point. Further research is being carried out on methods to tailor displays for low vision users.

One point to emphasize with regards to the reading test is that, in this pilot, the text is being presented on the display in a print-like format—static, black-on-white and uniformly spread across the display. Clearly, if the aim was to enhance reading speed, then far more could be done to make better use of the additional capabilities of an electronic platform. Indeed, this will be the next stage in the development process. However, even this basic display presentation method has its advantages over paper. A few participants described how these displays helped them read by removing background distractions and helping them concentrate on the text. The fact that displays are luminous would, no doubt, also have been of assistance. This would be a factor in the improvement seen in reading acuity from the displays. One reason the smartphone display may not have seen an overall improvement in reading acuity is that text size between 0.7 and 0.5 logMAR were included in the results, but were not perfectly rendered on the display.

The procedure used did not include practice time before the participants were timed to read sentences, which likely contributed to the average reading speed from both displays being slower than from paper. One participant stated, “If I got used to the headset I would prefer it”, citing their experience of becoming accustomed to their e-reader.

Although the text was rendered to subtend the same visual angle for each medium, our perception of size is related to our perception of the distance of the object. Viewing a virtual display confuses our perception of distance, particularly in the case of the smart glasses. One participant pointed out that when trying to read small print, established practice is to bring the print closer to the eye (and thus increase its visual angle). However, as the visual angle of the smart glasses display

is fixed, viewing the display on top of a distant wall makes it appear large, but moving closer to the background wall makes the display appear to shrink.

This highlights another important difference between the media that a different participant highlighted: When using paper, the participants retained control over the page and could move, tilt or rotate it as they pleased (whilst keeping it at a distance of 25 cm); but the HMD is held in a fixed position with respect to the head. This characteristic of the HMD may have contributed to the slower average reading speed from the displays, with one participant saying the smart glasses moved around a lot (with respect to the background) and another saying that the smartphone display was too close and they “wanted to take a step back”. However, we believe having the display in a fixed position is only a disadvantage if it is not tailored to the vision of the user. Further work is planned to investigate the use of the extent of visibility test for tailoring of the display.

The smart glasses rated more highly than the smartphone headset in terms of both comfort and how comfortable they would feel to wear the headset at home, but both were on the positive side of the rating scale. Several participants additionally commented that they would not hesitate to wear the smart glasses outside too. Their comments on prospective use, such as reading the news, labels on packets and buttons on the cooker, give insight into their challenges. A comment by a participant sums up an important factor in terms of prospective use: “If it made me see more, I’d use it!”

This pilot study suggests that HMDs are acceptable to those with AMD and that this pathology usually leaves a high proportion of the screen visible for viewing and reading from. A patient trial is planned to verify these claims amongst a wider patient population. Further research is required to investigate the tailoring of displays in light of data on extent of visibility. In addition, the inherent benefits of HMDs for improving reading performance need to be exploited as the technology matures and demand continues to grow.

**Acknowledgments** The authors would like to thank the following for their assistance: Peter Aspinall, Tariq Aslam, Ian Murray, Chi Can, Javid Khan, Ellie Brown, Cirta Tooth, Wolfgang Radner, Stephen Hicks, Douglas Young, Alex Bailey, Hannah Moshtael and the Macular Society. Supported by grants from Edinburgh & Lothians Health Foundation and the Engineering & Physical Sciences Research Council.

## References

1. Technavio Analytic Forecast: Global Smart Glasses Market for Augmented Reality 2015–2019 (2015)
2. Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G.P., Mariotti, S.P.: Global data on visual impairment in the year 2002. Bull. World Health Organ. **82**, 844–851 (2004)
3. Williams, R.A., Brody, B.L., Thomas, R.G., Kaplan, R.M., Brown, S.I.: The psychosocial impact of macular degeneration. Arch. Ophthalmol. **116**, 514–520 (1998)

4. Hazel, C.A., Petre, K.L., Armstrong, R.A., Benson, M.T., Frost, N.A.: Visual function and subjective quality of life compared in subjects with acquired macular disease. *Invest. Ophthalmol. Vis. Sci.* **41**, 1309–1315 (2000)
5. Massof, R.W., Rickman, D.L.: Obstacles encountered in the development of the low vision enhancement system. *Optom. Vis. Sci.* **69**, 32–41 (1992)
6. Weckerle, P., Trauzettel-Klosinski, S., Kamin, G., Zrenner, E.: Task performance with the Low Vision Enhancement System (LVES). *Vis. Impair. Res.* **2**, 155–162 (2000)
7. Culham, L.E., Chabra, A., Rubin, G.S.: Clinical performance of electronic, head-mounted, low-vision devices. *Ophthal. Physiol. Opt.* **24**, 281–290 (2004)
8. Werblin, F., Palanker, D.: Restoring vision to the blind: advancements in vision aids for the visually impaired. *Transl. Vis. Sci. Technol.* **3**, 54–62 (2014)
9. Hwang, A., Peli, E.: An augmented-reality edge enhancement application for google glass. *Optom. Vis. Sci.* **91**, 1021–1030 (2014)
10. van Rheede, J.J., Wilson, I.R., Qian, R.I., Downes, S.M., Kennard, C., Hicks, S.L.: Improving mobility performance in low vision with a distance-based representation of the visual scene. *Investig. Ophthalmol. Vis. Sci.* **56**, 4802 (2015)
11. Martin-Gonzalez, A., Kotliar, K., Rios-Martinez, J., Lanzl, I., Navab, N.: Mediated-reality magnification for macular degeneration rehabilitation. *J. Mod. Opt.* **61**, 1400–1408 (2014)
12. Moshtael, H., Aslam, T., Underwood, I., Dhillon, B.: High tech aids low vision: a review of image processing for the visually impaired. *Transl. Vis. Sci. Technol.* **4**, 6 (2015)
13. Radner, W., Willinger, U., Obermayer, W., Mudrich, C., Velikay-Parel, M., Eisenwort, B.: A new reading chart for simultaneous determination of reading vision and reading speed. *Klin. Monbl. Augenheilkd.* **213**, 174–181 (1998)
14. Orr, P., Rentz, A.M., Margolis, M.K., Revicki, D.A., Dolan, C.M., Colman, S., Fine, J.T., Bressler, N.M.: Validation of the national eye institute Visual Function Questionnaire-25 (NEI VFQ-25) in age-related macular degeneration. *Investig. Ophthalmol. Vis. Sci.* **52**, 3354–3359 (2011)

# **Method of Infrared Thermography for Earlier Diagnostics of Gastric Colorectal and Cervical Cancer**

**B. Dekel, A. Zilberman, N. Blaunstein, Y. Cohen, M.B. Sergeev,  
L.L. Varlamova and G.S. Polishchuk**

**Abstract** In this work we present a novel non-invasive method and the corresponding devices to diagnose ansinternal anomaly, that is, various kinds of intrinsic cancer, in a living subject by sending a passively occurring middle-infrared (MIR) radiation signal associated with the abnormality and inside an orifice of the Diagnostics includes detection and identification of the abnormality. A device or instrument is used either to bring a sensor into the orifice (in vivo diagnosis) or to transmit the MIR signal to the device or instrument located outside of the orifice (in vitro diagnosis). The example of instrument includes a prior art endoscope or gastroscope. The corresponding test results are presented as a proof of the proposed methodology of earlier diagnostics of internal cancerous structures.

**Keywords** Non-invasive method • Infrared thermography • Earlier diagnostics • Gastric colorectal • Cervical cancer

## **1 The Field and the Background of the Research**

During the recent decades it was stated a few common intrinsic cancers associated with orifices and the current art methods of diagnosis. Various embodiments were configured for detection, imaging and identification of colon cancer, cervical cancer, lung cancer, cancer of the esophagus, and stomach cancer, as well as of

---

B. Dekel · N. Blaunstein (✉)  
Scientific Center “Ruppin”, Netanya, Israel  
e-mail: nathan.blaunstein@hotmail.com

A. Zilberman · Y. Cohen  
PIMS Co., Beer Sheva, Israel

M.B. Sergeev  
ITMO University, Saint-Petersburg, Russia

L.L. Varlamova · G.S. Polishchuk  
LOMO, Saint-Petersburg, Russia

different kinds of internal tumors, lesions and other inner cancers by combined analyses of visible and infrared (IR) optical signals based on integral and spectral regimes for detection and imaging leading to earlier diagnosis and treatment of potentially dangerous conditions.

*Gastric cancer* is the seventh most frequent cause of cancer mortality over the world. The main screening methods for gastric cancer are:

*Upper endoscopy imaging*—A small visible spectral-range camera and light source are attached to a flexible guide and inserted through the throat and into the stomach of the patient. The resulting images are examined by a doctor and, if abnormalities are noted, tissue samples can be taken. This technique undergoes new improvements, like the ability to enable a zoom of magnification helping to identify the detailed surface structure [1]. Yet, this technique is a subjective and depends on the experience of the operator.

*Barium upper gastrointestinal radiographs*—For this test, people drink a barium-containing solution that coats the lining of the esophagus, stomach, and first portion of the small intestine, then, the radiologist takes multiple X-ray pictures. The accuracy of this test is not high, accomplished with other test can improve reliability [2].

*Endoscopic ultrasound*—A transducer probe placed into the stomach through the mouth or nose uses sound waves to produce images of internal organs. The transducer emits sound waves and detects the echoes bounced off internal organs. It is used to estimate how far cancer has spread into the wall of the stomach, to nearby tissues, and to nearby lymph nodes. To date endoscopic ultrasound is the most accurate imaging technique for staging depth of tumor invasion, not for preliminary detection [3].

*Computed tomography (CT)*—The CT scan is an X-ray procedure that produces detailed cross-sectional images of the stomach. Good sensitivity and accuracy are achieved with this non invasive method [4]. Main drawback is the X-ray radiation.

*Positron emission tomography (PET)*—In this test, radioactive glucose (sugar) is injected into the patient's vein. Because cancers use sugar much faster than normal tissues, the cancerous tissue takes up the radioactive material. A scanner can spot the radioactive deposits. This test, which is still being studied, is useful for spotting cancer that has spread beyond the stomach and can't be removed by surgery. It may be a very useful test for staging the cancer [5].

*Magnetic resonance imaging (MRI)*—MRI scans use radio waves and strong magnets. The energy from the radio waves is absorbed and then released in a pattern formed by the type of tissue and by certain diseases. A computer translates the pattern of radio waves given off by the tissues into a very detailed image of parts of the body. At present MRI appears to perform well in evaluating the local and distant extents of cancer but less well at detecting unsuspected primary tumors [6].

*Endoscopic autofluorescence spectroscopy*—A new technique based on emitting UV light for the excitation of tissue autofluorescence via the endoscope. Endogenous fluorescence spectra emitted by the tissue is collected with a fiberoptic probe and analyzed with a spectrograph. Yet results are not satisfactory [7].

*Colorectal cancer* is the third most common malignant neoplasm worldwide. The following current methods are used for colorectal screening [8–12].

*Fecal occult blood test*—The presence of hidden blood is detected in the stool. Blood in the stool that is not visible is often the first warning sign that a person has a colorectal disorder. The disadvantages of this method are that it detects blood in stool, but not its cause, and false-positive and false-negative results are common. Thus a more sensitive and precise test is needed.

*Flexible sigmoidoscopy and endoscopy*—These techniques are similar to upper endoscopy except that the endoscope is called a sigmoidoscope or colonoscope and inserted in the rectum rather than the throat. These techniques can discover from 50 to 65 % of polyps and are subject to all of the limitations of upper endoscopy.

*Virtual colonoscopy (CT Colonoscopy)*—Refers to examination of computer-generated images of the colon from data obtained by CT or MRI machines. The performance of this non-invasive method depends heavily on the size of the lesion. It can miss polyps smaller than 10 mm and generally suffers from the limitations of CT and MRI imaging mentioned above.

*DNA mutation of the stool*—This new non-invasive method is based on the detection of mutations in faecal DNA. At present the cost of this technique is high and sensitivity results are the same as colonoscopy.

*Barium enema*—Flow of barium is monitored on a X-ray fluorescence screen. This method has a low rate of detection even of large adenomas, but the technique is valuable in cases in which the colonoscopy does not reach the lesion [13].

*Cervical cancer* is cancer of the uterine cervix, the portion of the uterus attached to the top of the vagina. Ninety percent of cervical cancers arise from flattened or “squamous” cells converting the cervix. Most of the remaining 10 % arise from the glandular, mucus-secreting cells of the cervical canal leading into uterus. This cancer is the second most common cancer in the women worldwide. The following methods are usually used for cervical screening and detection [8–10].

*Pap smear*—This screening examination is obtained by collecting a sample of cells from the cervix with a wooden or plastic spatula and brush. Specimens are placed on glass slides and examined by a special pathologist or cytologist. If abnormalities are found, women are typically asked to return for colposcopy. The quality of the pap smear technique can be compromised by inflammatory exudate or failure to sample the transformation zone. As a result, a relatively high false-negative rate of 20 % pap smears might cause failure to diagnose pre-invasive disease.

*Colposcopy*—This technique uses a magnifying lens to view the surface of the cervix under white and green light after a mild vinegar solution is applied. If pathologic areas are seen, a biopsy is taken. This method is not performed in real time and has the disadvantages of other forms of visible light endoscopy as was mentioned above. Particular, visible light endoscopy is subjective and depends on physician experience alertness.

It should be stated that none of above mentioned techniques of detection of all kinds of cancers are capable of positively identifying tumors. Therefore according to current art distinguishing tumors from other benign or pathological conditions

requires biopsy. However, biopsies have many obvious disadvantages which we briefly reflect: firstly a biopsy requires intrusive removal of tissue that can be painful and expensive. Particularly, in internal cavities and more particularly, in the stomach and intestines, biopsies run a high risk of serious complications. These simplifications can lead to very painful conditions, including ulcers, they can force limiting diet or activity of a patient for significant periods of time and complications may even require treatment and drastic interventions (for instance surgery). Only a very limited number of sites can be biopsied in one session. Secondly, biopsy samples must be stored and transported to a laboratory for expert analysis. Storage and transportation increase the cost, increase the possibility that samples will be mishandled, destroyed or lost, and also cause a significant time delay in receiving results. This time delay means that examination follow up requires bringing the patient back to the doctor for a separate session. This increases the inconvenience to the patient, the cost and the risk that contact will be lost or the disease will precede to a point of being untreatable. Furthermore, the waiting period causes significant anxiety to the patient. Finally, interpretation of biopsies is usually by microscopic analysis, which results in qualitative subjective results that are not well suited to consistent interpretation.

Therefore, in medical diagnosis, there is great interest improved sensitivity, safe non-operative detection technologies capable of revealing internal cancers in their earlier stage and also in improved techniques for identification to differentiate between cancer, benign conditions and other pathologies of internal tissue.

The present work relates to a novel method and the corresponding devices to detect and identify pathologies inside orifices of a living subject and more specifically to a method and devices to detect and identify gastric, colorectal and cervical cancer, based on near infrared (NIR) and middle infrared (MIR) devices and instruments.

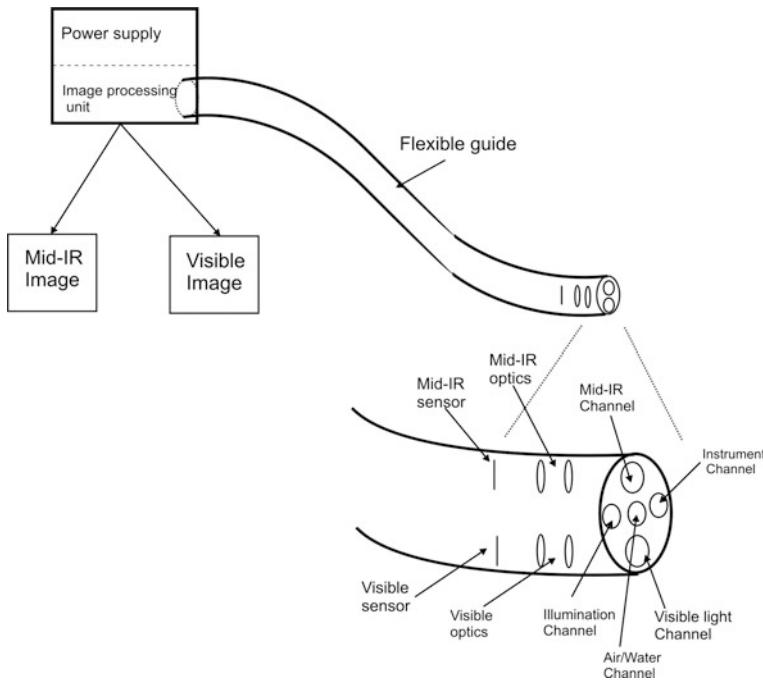
## 2 The Proposed Methodology and the Device

The proposed method is based on the detection of malignant tissues related to the gastric, colorectal and cervical cancers based on two techniques [14]:

1. Infrared imaging method-NIR and MIR.
2. NIR and MIR spectral method

Optical methods of imaging were coming from detection and diagnosis of skin cancer. We do not enter into this subject because it is out of scope of our work, referring the reader to excellent papers [15–17].

As for spectral regime, using optical radiated signal ranged at the wide frequency bands, it based on perceiving reflected light in the visible and near-infrared (NIR) bands. Identification by use of this regime of specific abnormalities is based on information about the corresponding “signature” of radiation associated with the



**Fig. 1** Schematical presentation of the flexible guide

corresponding anomaly measured in the frequency domain. Recently, middle-infrared (MIR) spectral methods have also been used to improve an accuracy and reproducibility of biopsy evaluation for the gastric and stomach cancers [18–21].

In our case of interest, the image capturing is based on usage of an endoscope consisting of a NIR or MIR camera based on micro-bolometric or pyroelectric detector array mounted on the distal tip of the flexible tube (as shown in Fig. 1) and a conventional visible camera. In addition to capturing image frames in the visible and Mid-IR spectral regions, light source, air, water and small instruments can be used to take samples of suspicious tissues through the endoscope. This device is inserted through the mouth or anus the same method as it is done by the visible endoscopes at present. Schematic drawing of the endoscope is shown in Fig. 1.

The flexible guide consist of five channels:

- (1) Mid-IR channel, which includes a window in the distal tip and a series of optical elements in a wide field configuration, all are transparent in the Mid-IR spectral range (like Zinc Selenide, Silver halide and others), and a microbolometric CCD.
- (2) Visible channel the same as the Mid-IR channel except that the materials of the components are transparent in the visible spectral range.
- (3) Air/water channel delivers air, water or to suction intraluminal contents.

- (4) Illumination channel provides illumination in the visible spectral range to the suspected tissue.
- (5) Instrument channel provides the physician tools to take samples from suspected tissues.

The power supply unit is responsible to maintain the Mid-IR detector, the visible detector and the illumination unit in operational mode. The image processing unit delivers two video images captured by the two sensors, a Mid-IR thermal image and a visible image. This unit is carrying out image processing algorithms and performing the contrast algorithm to improve the probability of detection.

There are two methods of tissue surface and sub-surface anomalies detection and identification. The first method is called the *integral regime* of tissue detection using information about the corresponding gradient of temperature between the normal tissue and the cancerous tissue. The second method is called the *spectral regime* of tissue detection using information about the spectral lines of irradiated field from cancerous tissue, that is, about the corresponding “signature” of the anomaly in the frequency domain.

The broad spectrum from NIR- to MIR-band is used for diagnostics and identification of the tissue anomaly. This method is based on the integral and spectral thermograph analysis of infrared flows radiated by the living tissue surface, the cancerous and the regular tissue. It consists:

- (1) Measure of the space distribution of the integral flow from the object at the range of wavelength from 3–5  $\mu\text{m}$  (NIR-band) to 8–12  $\mu\text{m}$  (MIR infrared), most often from the open surface of the body.
- (2) Spectral regime based on the following parameter such as spectral density of emitting radiation ( $dR/d\lambda$ ) from human body as a black body with temperature  $T_0 \approx 36.6^\circ\text{C}$  using visual spectrum from 200 to 900 nm and infrared spectral range of 3.0–20  $\mu\text{m}$ . The normalized heat flow difference (or contrast) was measured in the anomalous zones in the chosen wavelengths with the known narrow waveband  $\Delta\lambda_i$  of measurements. The counting of the mean spectral density of the measured heat flows in each band of measuring according the formula:

$$S_{\lambda i} = R_{\lambda i} / \Delta\lambda_i, \quad (1)$$

where

$S_{\lambda i}$  is the mean spectral density of the heat flow for the chosen;

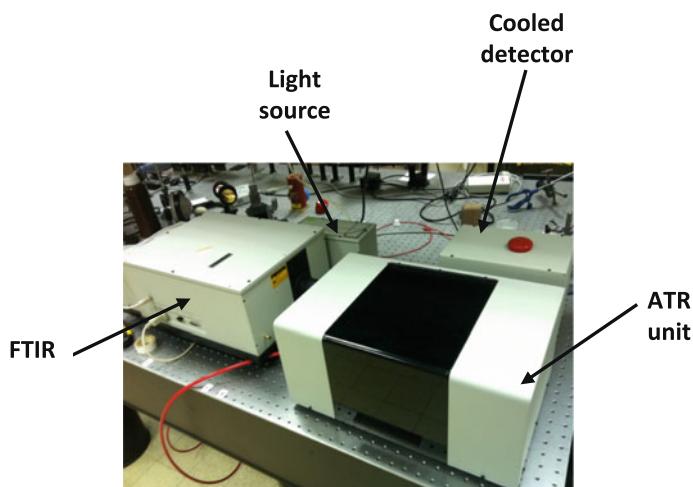
$\lambda_i$  is the bandwidth (ith wavelength);

$R_{\lambda i}$  is the measured value of the heat flow in the chosen  $\lambda_i$  band;

$\Delta\lambda_i$  is the spectral width of the chosen ith band.

### 3 Applications of NIR and MIR Techniques and Devices

For identification of gynecologic cancer cells during a surgical procedure, an in vitro system based on MIR-ATR spectroscopy was designed for the purpose of examination of the adequacy of MID-IR spectroscopy, in conjugation with ATR (see Fig. 2) for the detection of gynecological malignancies in real time during surgery. The second device with the corresponding ATR-MIR spectroscopy instruments for identification of the colonoscopic and gastric cancer cells during surgical procedure was developed recently (see Fig. 3). Its purpose was to examine

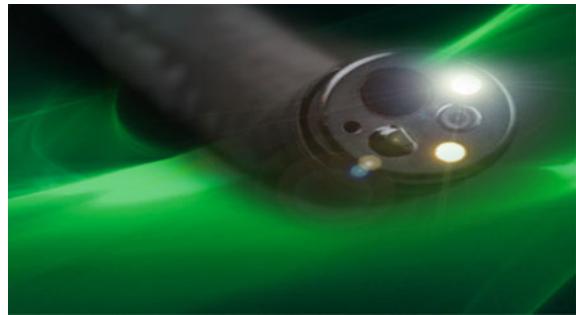


**Fig. 2** An in-vitro operated MIR-ATR system

**Fig. 3** The open view on an in-vitro operated MIR-ATR system



**Fig. 4** The view of MIR-ATR probe



the adequacy of middle infrared spectroscopy in conjugation with ATR, for the detection of colonoscopic and gastroscopic malignancies in real time during surgery. Recently, this device in a series of clinical tests at the medical center in Jerusalem. During this stage the system and histopathology data were taken in parallel and the correlation was computed between the two measurements. About 70 samples were taken up to date. In the next stage our system will classify the biopsies.

Then, the possibility to check *in vivo* all cancerous cells during surgical procedure was investigating and the corresponding hollow fiber and a MIR-ATR probe were performed based on the technology presented and described in Fig. 1. The corresponding product is shown in Fig. 4.

Such a probe allows us to identify in real time whether tissue is malignant. The technique involves inserting a hollow optical fiber in the working channel of the endoscope and detecting the signal from the ATR probe.

Moreover, the developed probe based on hollow fiber technology allows to construct an *in vivo* system for enhancing Colposcopy procedure, using the spectral properties of cervical cancer in the NIR spectral range. The main goal of usage of such a MIR-ATR probe was to improve the detection precision in the Colposcopy procedure by using the unique properties of the malignant tissues in the combined visual optic system (VIS) and NIR spectral ranges that now exists separately.

## 4 Summary

Using Near- and Mid-IR techniques for detecting and identification of gastric, colorectal and cervical cancers is different with respect to other methods in that the conventional methods are based on the subjective inspection in the visible spectral range of the physician, where in our method we are based on the overheated characteristic of cancerous tissues and detecting these anomalies with a NIR or MIR camera.

In our work, based on different kinds of instruments and devices, the detection and identification of cancers using the combination of the visible, NIR and spectral

regimes in combination with the contrast, is different with respect to other methods in that the preliminary classification of anomaly is done, using IR optical band in the range of 3–12 μm.

We developed a new endoscope based on an array of infrared and visual optic micro detectors operating in the waveband range of 3–12 μm working both in integral and spectral regimes based on the sign and amplitude of the contrast of cancerous anomalies with respect to those for normal living tissues.

## References

1. Sugano, K., Sato, K., Yao, K.: New diagnostic approaches for early detection of gastric cancer. *Digest. Dis.* **22**(4), 327–333 (2004)
2. Bender, G.N., Makuch, R.S.: Double-contrast barium examination of the upper gastrointestinal tract with non-endoscopic biopsy: findings in 100 patients. *Radiology* **202** (2), 355–359 (1997)
3. Messmann, H., Schlottmann, K.: Role of endoscopy in the staging of esophageal and gastric cancer. In: *Seminars in Surgical Oncology*, vol. 20, no. 2, pp. 78–81 (2001)
4. Bhadari, S., Shim, C.S., Kim, J.H. et al.: Usefulness of three-dimensional, multidetector row CT (virtual gastroscopy and multiplanar reconstruction) in the evaluation of gastric cancer: a comparison with conventional endoscopy, EUS, and histopathology. In: *Gastrointestinal Endoscopy*, vol. 59, no. 6, pp. 619–626 (2004)
5. Yoshioka, T., Yamaguchi, K., Kubota, K. et al.: Evaluation of 18F-FDG PET in patients with a metastatic, or recurrent gastric cancer. *J. Nucl. Med.* **44**(5), 690–699 (2003)
6. Motohara, T., Semelka, R.C.: MRI in staging of gastric cancer. *Abdomin. Imaging* **27**(4), 376–383 (2002)
7. Mayinger, B., Jordan, M., Horbach, T. et al.: Evaluation of in vivo endoscopic autofluorescence spectroscopy in gastric cancer. *Gastrointest. Endosc.* **59**(2), 191–198 (2004)
8. Shike, M., Winawer, S.J., Greenwald, P.H. et al.: Primary prevention of colorectal cancer. The WHO collaborating center for the prevention of colorectal cancer. In: *Bull World Health Organ*, vol. 68, no. 3, pp. 377–385 (1990)
9. Judith, M., Walsh, E., Terdiman, J.P.: Colorectal cancer screening: scientific review. *JAMA* **289**(2003), 1288–1296 (2003)
10. Imperiale, T.F., Wagner, D.R., Lin, C.Y. et al.: Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *New Engl. J. Med.* **343**(3), 169–174 (2000)
11. Ferrucci, J.T.: Colon cancer screening with virtual colonoscopy: promise, polyps, politics. *AJR Am. J. Roentgenol* **177**(5), 975–988 (2001)
12. Traverso, G., Shuber, A., Olsson, L. et al.: Detection of proximal colorectal cancers through analysis of faecal DNA. *Lancet* **359**(9304), 403–404 (2002)
13. Winawer, S.J., Stewart, E.T., Zauber, A.G. et al.: A comparison of colonoscopy and double-contrast barium enema for surveillance after polypectomy. National Polyp Study Work Group. *New Engl. J. Med.* **342**(24), 1766–1772 (2000)
14. Dekel, B.Z., Blaunstein, N., Zilberman, A.: Method of Infrared thermography for earlier diagnosis of gastric colorectal and cervical cancer. US Patent: US 8,774,902 B2, July 2014, 11 pages (2014)
15. Gniadesca, M., Wulf, H.C., Nymark, N., et al.: Diagnosis of basal cell carcinoma by Roman spectroscopy. *J. Roman Spectrosc.* **28**(1997), 125–129 (1997)

16. Brooks, A., Afanasyeva, N.I., Makchine, V., et al.: New method of investigations of normal human skin surfaces *in vivo* using fiber-optic evanescent wave Fourier Transform infrared spectroscopy (FEW-FTIR). *J. Surf. Interf. Anal.* **27**(1999), 221–229 (1999)
17. Yang, Y., Sule-Suso, J., Sockalingum, G.D. et al.: Study of tumor cell invasion by Fourier Transform infrared microspectroscopy. In: *J. Biopolymers* **78**, 311–317 (2005)
18. Fujioka, N., Morimoto, Yu., Arai, T., et al.: Discrimination between normal and malignant human gastric tissues by Fourier transform infrared spectroscopy. *J. Cancer Detect. Prevent.* **28**(2004), 32–36 (2004)
19. Wang, H.P., Wang, H.C., Huang, Y.J.: Microscopic FTIR studies of lung cancer cells in pleural fluid. *J. Sci. Total Environ.* **204**, 283–287 (1997)
20. Christodoulou, S., Chrysohoon, Ch., Panagiotakos, D.B., et al.: Temperature differences are associated with malignancy on lung lesions: a clinical study. *Bio Med. Central* **2003**, 1–5 (2003)
21. Christodoulou, S., Paeskevas, E., Panagiotakos, D.B., et al.: Thermal heterogeneity constitutes a marker for the detection of malignant gastric lesions *in vivo*. *J. Clin. Gastroenterol.* **2003**, 215–218 (2003)

# Tijuana's Sustainability for Healthcare Measurement Using Fuzzy Systems

**Bogart Yail Márquez, Arnulfo Alamilis,  
Jose Sergio Magdaleno-Palencia, Karina Romo, Felma González  
and Sergio Méndez-Mota**

**Abstract** The proposed methodology is focus as an alternative to analyze and describe the most accurate social phenomena according our reality using different computational mathematical theories, which are not used conventionally in social sciences applications and this is a new approach to create new computer's simulation architectures.

**Keywords** Fuzzy logic • Natural phenomena • Data mining • Sustainable development

## 1 Introduction

This project is oriented to study the main characteristics related to border's cities sustainability, defining relations, actions and rules that exist in real world and these are difficult to represent statistically; the methodologies are based on complex social

---

B.Y. Márquez (✉) · A. Alamilis · J.S. Magdaleno-Palencia · K. Romo · F. González · S. Méndez-Mota  
Baja California Autonomous University, Calzada Universidad 14418,  
Tijuana, Baja California, Mexico  
e-mail: bogart@tectijuana.edu.mx

A. Alamilis  
e-mail: alanan@tectijuana.edu.mx

J.S. Magdaleno-Palencia  
e-mail: jmadaleno@tectijuana.edu.mx

K. Romo  
e-mail: karinaromeroalvarado@gmail.com

F. González  
e-mail: felmagonzalez@gmail.com

S. Méndez-Mota  
e-mail: smonix359@gmail.com

simulations, these need to be confirmed by specific necessities and main relation from sustainability.

This research is motivated by the existing problems in border's cities and specifically in the North of Mexico, thus since in recent decades this city has grown faster than in the rest of the country and this affect directly in the cities sustainability. This trend, together with patterns of land occupation where the efficiency in land use has not been promoted actively, it has generated urban fragmentation areas with high population density, a high percentage of interurban unbuilt lots and average population densities as is the case of Ciudad Juarez where the population is 1,321,004 and Tijuana with a population of 1,559,683, these are the most populated cities in the State of Baja California, Mexico [10].

The border's cities are closer to the United States, and that aspect influences directly to trade, migration, the source of employment and foreign exchange earnings for this area. previous social research in this area tell us that the new sub centers are the result of industrial expansion forces, all of this around of international crossings borders and distant room building new areas that demanded new trade centers distance, and also it defines that expansion transformed the urban structure of a city; it is necessary analyze different morphologies [7].

Some studies done using zonal based models were elaborated for Third World cities, such studies pretended explain the patterns in land use, with the social group identity, this studies were looking for to establish hegemonic behavior within the urban space. Different urban development plans, the expansion of the urban sprawl, this expansion is done by a process of continuing land occupation, in the reality it is very different because this occupation is very irregular and it is done in isolated zones that are not connected between them, or in the best connected by traditional ways, such as gaps or dirt roads [11].

## 2 Identification of the National Problem to Be Addressed

The situation of the Mexico's North Border has meant a series of problems, many of them are perceived in daily life. The borders represent to many people the opportunity to get better jobs, more salary and as a result of this in better living conditions; it represent the access to the world's largest market for domestic companies; also for criminal organization, the North Borders is the right place to negotiate drugs and also to get guns to their criminal activities [12].

Internal migration was routed toward the poles whose economic activity was more intense. In the North of the country, the population increase and it is because exists more jobs, which it influences to the population mobility. In three cities the migration increase considerably: Tijuana, Mexicali and Ciudad Juarez, this city are the most representatives of this phenomena, our focus of study are in these cities [9].

The study will address the issue of "movements and human settlements" and a model will be generated considering historical tendencies of the increasing urban sprawls, and the fundamental of economic basic of the population distribution, the

economic activities in the territory will be associated to this study. In addition, this study will increase the accuracy of the bearings and predominant directions associated with the dispersion of the urban area at different times, so they can be settle better if they are opened and be related and they are opened to new areas to urbanization, with the development of large public and private projects, and of course the attraction that they tend to generate [8].

In the other hand, the widespread presentation of the physical expansion is not provided for a formal document that shows and gives information of its behavior, sustained in a theoretical document, this research is looking to fill this document, and valorizing the importance of the different phenomenal that have been used in the city delimitation. In this regard, the development of this research will be based on the theory of Complex Social Systems.

## Objectives

Our objective is to analyze and simulate the sustainable development in border cities more important of the country. Specifically according the following:

- Create a sustainable simulation model that helps to experiment and predict social problems avoiding costs and possible risks.
- Analyze and predict human settlements in border cities
- Analyze and predict water consumption in border cities and analyze the convergent problematic
- Analyze and predict population growth of border cities.
- Analyze and predict the generation of employment in border cities.

## Scientific Methodology

The concept of Fuzzy Logic, was conceived by Lotfi Zadeh a professor at the University of California at Berkeley, who disagreed with the classical sets (crisp sets), which allows for only two options; membership or not an item to the all presented as a way of processing information about allowing partial memberships joint as opposed to the classic called Fuzzy Sets. If we make a comparison between the classical and fuzzy logic we can say that classical logic provides a logical parameters between true or false, that is, using binary combinations of 0 and 1, 0 if false and 1 if true. Now if we take into account that the fuzzy logic which introduces a function that expresses the degree of membership of an attribute or variable to a linguistic variable taking the values between 0 and 1, this is called fuzzy set and can be expressed mathematically:

$$A = \{x / \mu_A(x) \forall x \in X\} \quad (2)$$

This work is motivated by the need to establish a methodology for the study of complex social systems, because it is going to has a difference with previous studies, this methodology is intended to encompass issues as adaptive and emergent, such situation in which the conventional analysis is not enough to describe the complexity of real social phenomenal and the social actors. The proposed methodology in

general describes the use of several computational techniques and interdisciplinary theories. This continuous consensus growth must be capable to describe all social life aspects as well as serve as a common language in which the different theories can be proven.

Our project has the main objective to create a sustainable system model, for urban growth, to achieve this reproduction, we will use computer science techniques, depending on the economic social level [4].

Dynamic systems, it will allow us to represent all the elements and relations of the structure sustainable system and the evolution according time. Also the system will allow obtain the mathematical equations of the model at macro level [1, 3].

Multi-agent systems (MAS), consists of autonomous agents that work together to solve problems, characterized in that each agent has information.

Incomplete ability to solve the problem [2], there is no global control system, data are decentralized and the computation is asynchronous [14]. The agents dynamically decide which task is going to be done. Using geographic information of each city divided by AGEBS, where each AGEB delimit a part or the whole of a town of 2500 inhabitants or more, each AGEB representing an agent and contains a quantitative information about employment, education, income, and infrastructure that has a location.

Cellular automata (AC), this technique will be used to simulate human settlements in space, it is a useful tool for modeling a geographic system. AC meanwhile structures are ideal for direct and discrete representation of complex behaviors at the cellular level [5], multicellular and population. The modeling of the project is based on methods and mathematical expressions [13], which theoretically represent the behavior of land use of our model [6].

### 3 Methodology

Since the appearance of the unmanned aerial vehicles in the early days these were develop by the United States for surveillance reconnaissance purposes in the World War 1 and World War 2 and then they were widely used in the 20th century between the 60s, 80s and 90s then several countries begin using the unmanned aerial vehicles and now they are worldwide used by military and civilian, the main purpose was the surveillance of individuals and aerial reconnaissance but now days has been incorporated as a means of recreation and now they're taking further steps not just as a recreation but now as a tool for several purposes such as photography, Acrobatic filmmaking, aerial footage reconnaissance, real estate inspections, search and rescue operations and some other uses, here we are focusing on mapping for geographical purposes. There are several types of vehicles some are called fix wing and other are called quadcopters, in both cases they require a camera for the taking of the aerial imagery, this is called aerial photogrammetry, the image taken by the camera have to be process by a software for the purpose of making the mapping of the area, building or object. By using a digital camera we have several advantages

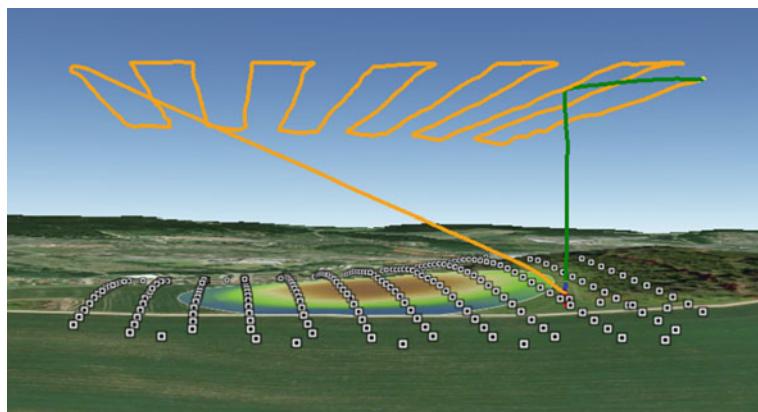
these include easy of used, handy, imagery in digital format that is ready to be use and depending of the type of camera and size this cameras can be put into a several type of aerial vehicles depending on their capabilities and some of them can be equipped with a global positioning system GPS. Predicting the population growth according to the current inertia, to solve the problems that generate a pattern without a sustainable urban development, that only resolve land necessities in a expansive way and without considering ecological care required by border cities. Promote reflection of environment care in consequence with the exponential settlements land use.

Table 1 shows the category of UAV defined by Unmanned Vehicle Systems-International (UVS) (Ahmad, A. Digital Mapping using Low Altitude UAV, 2011).

The green line is the point of origin, the yellow line is the path that the UAV will take and the small square dots are the images capture during the flight path, once we have the information then we can continue with the process of forming a chain of images as the following image shows (Fig. 1).

**Table 1** Category of UAV

Category name	Mass (kg)	Range (km)	Flying altitude (m)	Endurance (h)
Micro	<5	<10	250	1
Mini	<250/30/150	<10	150/250/300	<2
Close range	25–150	10–30	3000	2–4
Medium range	50–250	30–70	3000	3–6
High alt. long endurance	>250	>70	>3000	>6



**Fig. 1** UAV

## References

1. Bar-yam, Y.: *Dynamics of Complex Systems*. Addison-Wesley, Reading (1992)
2. Bellifemine, F., Caire, G., Greenwood, D.: *Developing Multi-agent Systems with JADE*. John Wiley & Sons Ltd., West Sussex (2007)
3. Berne, E.: *The Structure and Dynamics of Organizations and Groups*. Freemantle Publishing, Beaconsfield (2001)
4. Byrne, D., Callaghan, G.: *Complexity Theory and the Social Sciences: The State of the Art*. Taylor and Francis (2013)
5. Coulhon, T., Mongin, P.: Social choice theory in the case of von Neumann-Morgenstern utilities. *Soc. Choice Welfare* **6**(3), 175–187 (1989)
6. Evans, D., Leao, S., Bishop, I.: Simulating urban growth in a developing nation's region using a cellular automata-based model. *J. Urban Plann. Dev.* © ASCE/Sept. 2004. doi:[10.1061/\(ASCE\)0733-9488\(2004\)130:3\(145\)](https://doi.org/10.1061/(ASCE)0733-9488(2004)130:3(145)) (2004)
7. Geanakoplos, J., Karatzas, I., Shubik, M., Sudderth, W.: Inflationary Equilibrium in a Stochastic Economy with Independent Agents (2009)
8. Glaeser, E., Gyourko, J., Saks, R.: Urban growth and housing supply. *J. Econ. Geogr.* **6**, 71–89 (2005)
9. Herrera, L.A., Pineda Jaimes, S.: Ciudad Juárez: Las sociedades de riesgo en la Frontera Norte de México. *FERMENTUM Revista Venezolana de Sociología y antropología* **17**, 422 (2007)
10. INEGI: *Instituto Nacional de Estadística y Geografía* (2010)
11. Márquez, B., Castanon-Puga, M., Suarez, E.D., Magdaleno-Palencia, J.: Modeling the employment using distributed agencies and data mining. In: Zhu, M. (ed.) *Electrical Engineering and Control*, vol. 98, pp. 933–940. Springer, Berlin (2011)
12. Mendoza, J.E.: El mercado laboral en la frontera norte de México: estructuras y políticas de empleo. *Estudios Fronterizos* **11**(21) (2010)
13. Morgenstern, O.: The collaboration between oskar morgenstern and John von Neumann on the theory of games. *J. Econ. Lit.* **14**(3), 805–816 (1976)
14. Phan, D., Varenne, F.: Agent-based models and simulations in economics and social sciences: from conceptual exploration to distinct ways of experimenting. *J. Artif. Soc. Soc. Simul.* **13**(1), 5 (2010)

# **Intelligent System for Learning of Comfort Preferences to Help People with Mobility Limitations**

**Sandra López, Rosario Baltazar, Lenin Lemus Zuñiga,  
Miguel Ángel Casillas, Arnulfo Alanis, Miguel A. Mateo Pla,  
Víctor Zamudio and Guillermo Méndez**

**Abstract** It is essential to provide better living conditions for vulnerable sectors of society using technology and it is important to consider that the technology must be friendly with users, and even adapt to their needs and desires. We can see in current systems the user has to learn to use the devices or services, but with an intelligence system, the technology is a very effective way to determine the needs of users. In this research, we present the physical implementation of a system to assist users and patients in daily activities or duties. The system include a architecture of agents where a Deliberative agent learns from the interaction with the user, in this way the system detects thermal comfort preferences for give an automatic assistance. We propose an algorithm with a proactive stage and learning stage adapting a classification algorithm. We select the classification algorithm with the best performance using cross validation. The algorithms of pattern recognition was Back Propagation

---

S. López · R. Baltazar (✉) · M.Á. Casillas · V. Zamudio · G. Méndez  
Instituto Tecnológico de León, Guanajuato, Mexico  
e-mail: r.baltazar@ieee.org

S. López  
e-mail: jclsandra@gmail.com

M.Á. Casillas  
e-mail: miguel.casillas@gmail.com

V. Zamudio  
e-mail: vic.zamudio@gmail.com

G. Méndez  
e-mail: guillermomendez06@gmail.com

L.L. Zuñiga · M.A.M. Pla  
Universitat Politècnica de València, Valencia, Spain  
e-mail: lemus@upv.es

M.A.M. Pla  
e-mail: mimateo.pla@gmail.com

A. Alanis  
Instituto Tecnológico de Tijuana, Baja California, Mexico  
e-mail: alanis@tectijuana.edu.mx

neural network, Naïve Bayes, Minimum Distance and KNN (k near neighbor). Our motivation of this work was to help people with motor difficulties or people who use wheelchairs, for this reason it was essential to use a wireless controller and use a friendly interface. The system was implemented in a testbed at the Leon Institute of Technology in Guanajuato, Mexico, and include sensors of humidity and temperature, windows actuators, wireless agents and other devices. Experimental tests were performed with data collected during a time period and using use cases. The results were satisfactory because it was not only possible remote assistance by the user but it was possible to obtain user information to learn comfort preferences using vector features proposed and selecting the classification algorithm with better performance.

**Keywords** Assist in daily activities • Movement restrictions • Architecture of agents • Learning of thermal comfort • Pattern recognition • Classification algorithm

## 1 Introduction

The interaction of people with technology is more and more common, it is easy to see how computers, mobile devices, etc. are increasingly used, so we could say that are used naturally. However, in almost all this devices, the user has to use them continuously for being acquainted with them.

Unlike current computing systems where the user has to learn how to use the technology, an IE (intelligent environment) adapts its behavior to the users, even anticipating their needs, preferences, or habits. For this reason, the environment should learn how to react to the actions and needs of the users, and this should be achieved in an unobtrusive and transparent way. In order to provide personalized and adapted services, it is necessary to know the preferences and habits of users [1].

There are many works about learning of users, the researchers at Essex's iDorm lab have given prominence to the problem of learning. They are one of the most active groups in the area of fuzzy logic control on Ambient Intelligence. Their initial efforts (Hagras et al. in [2] and Doctor, Hagras, and Callaghan in [2]) were focused on developing an application that generated a set of fuzzy rules representing a user's patterns. Vainio et al. [3], also used fuzzy rules to represent habits of the user. In contrast to the approach followed in the iDorm project, these authors manually constructed the membership functions and used reinforcement learning to replace old rules.

The group that has been working on the MavHome and Casas projects is one of the most active groups in the area of sequence discovery. The first applications developed by this group were focused on building universal models, represented by means of Markov models, in order to predict either future locations or activities (Cook and Das [4], Heierman and Cook [5] and Rao and Cook [6]).

According to expose in [7], the use of ANNs has a limitation related to their black box nature; therefore their internal structure is not human-readable. Other authors (Stankovski and Tmkoczy [8]) have pointed out that the model generated by decision trees is easy to understand and suitable for human inspection.

Aztiria et al. [1], have started to use the strategy of combining different techniques in order to identify complex patterns in Intelligence Ambient environments.

In application for remote healthcare system, in [9] the users can monitor their healthy lifestyle through daily check-ups and make changes based on the prescription provided. In [10], the authors propose a system that has been designed to monitor elders which live alone and want to keep living independently. An important contribution in this work is related with the capability of the system to adapt its behaviour to that of the monitored elder. In [11, 12], the authors developed a framework where the agent agents adapt the environment depending on the user's emotions.

Other authors, as [13], have already developed a system that consists of two basic modules. At the lower level, Wireless Sensor Network (WSN) capture the sensor data based on the usage of house-hold appliances and store the data in the computer system for further data processing.

We present an architecture of agents that have the ability of monitoring data of the user. At the same time the user uses the system in a remote way to make his daily activities. When the system learn the patterns of the user this react in the environment for give a personalized assistance to the user. We decided use a neural network and other classification algorithms because they are used very sparingly in learning of thermal comfort preference.

One of our main motivations for this work was to provide independence, privacy and dignity to elderly persons or people with physical disabilities. Another important aspect that motivates us to make this work, was the fact that there is the problem of finding the right person for comfort. In [14], the author says that the task of finding the optimum values for the human body has the feeling of comfort for all people is virtually impossible, due to the great physiological and psychological differences of individuals. The best way to solve the problem of finding these values is to use the “Charter effective temperature” or “Letter of Comfort”, but this way is a general solution, this mean that it has no customization.

## 2 System Description

This section describes the development and implementation of an intelligent system for learning a user, that goes beyond simple assistance but based on their interaction with the system, is realized learning of behaviors that belong at user. To do this, first the architecture and properties of the agents were designed, developed and implemented in a room that is used continually. The room is in the Leon Institute of Technology in Guanajuato, Mexico, see Fig. 1.



**Fig. 1** Image of the room of test

## 2.1 Design of Architecture Agents

Agents sensors were developed to detect environment information, actuators agents to manipulate external elements, same that interact with the user, and to process and route data, agents, information managers were developed and a deliberative agent where were implemented artificial intelligence techniques, among others. The system consists of the following agents:

*User Agent:* It is considered the same user that interacts with the environment, as an agent of the system. Therefore, sends actions of control to the different actuators, this is realized through a mobile device using Bluetooth communication. This agent has the property of be deliberative.

*Agent User Interface:* Through this agent, the user agent can communicate with the system. This agent comprises a graphical interface implemented on a mobile device.

*Agent Gateway:* Its function is change a communication protocol in another.

*Deliberative Agent:* Responsible to request registers in a database of historical information and current information from the different sensors agents implemented in the environment. Based on this information the agent issues a decision processed by its intelligent algorithm, so that their response is sent to the various actuators corresponding.

*Management Information Agent (IMA):* Is responsible for sending the messages issued by the Deliberative Agent or User Agent, to the various sensors agents, actuators and viceversa. Management Agents information receives the message from the Deliberative Agent via an interrupt and store it in a queue, so it sends the

message to its proper destination in order of arrival (first in, first out), so does the response messages received from the sensors agents and actuators.

*Agents Sensors:* They are responsible for receiving the order of the Deliberative Agent or the Agent Middleware that is sent through agents Information Management, to take a measurement. Such measurement is returned to the agent that issued the request through the Information Management Agents.

*Agents Actuators:* They are responsible for receiving the order of the Deliberative Agent or User Agent that is send through the Information Management Agents to perform a certain action.

*Middleware Agent:* Responsible for receiving messages with different communication protocols to send data to the database agent. It also sends orders each certain time, to obtain information of the sensors via Agents Information Management, so that the measurement is returned to the Agent Middleware for it to send it to Database Agent.

*Database Agent:* Responsible for keeping records that are sent from the Middleware agent, saving the time and date of registration as well as the identifier and the registry value.

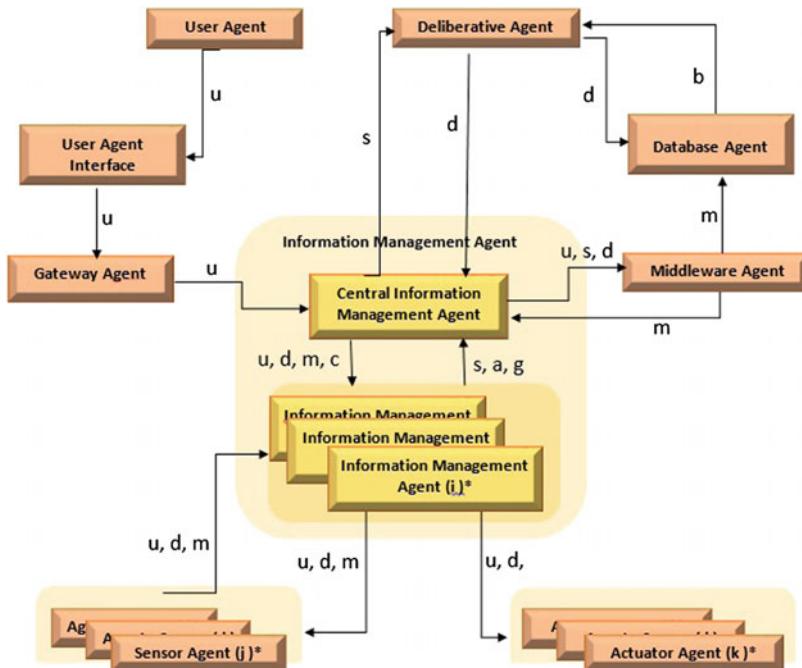
The Agents architecture designed can be seen in Fig. 2, where in this is also presented information flows system. The system consists of the following information flows: u (Information flow dependent of user agent), d (Information flow dependent of deliberative agent), m (Information flow dependent of middleware agent), b (Information flow dependent of database agent), c (Information flow dependent of management core information agent), g (Information flow dependent of information management agent), s (Information flow dependent of sensor agent), a (Information flow dependent of actuator agent).

Communication between the various IMA, is designed in form centralized, i.e. messages sent by the deliberative agent, user agent and agent middleware arrive the ACIM (Agent Central Information Management) to manage the different orders of others IMA and reach their destination.

## 2.2 *Detailed Description of the Hardware of the Agents*

We developed 12 agents for information management, 8 temperature sensor and humidity sensor agents, 4 agents window actuators, 1 Gateway agent, 1 interface agent. We use the hardware of the System in [15], in [16] and the middleware in [17].

It was selected the DHT22 temperature-humidity sensors for agents which uses the I2C protocol to transmit measured values. In order to communicate the sensor agent with IMA, an electronic board was developed with a MSP430 microcontroller from Texas Instrument, the microcontroller is programmed to perform 2 different functions (synchronization sensor with IMA and receive the order to send data of temperature-humidity). We developed Window actuators agents. This have the function of move sliding windows that have a dimension of 1.20 m.



\* $i, j \text{ y } k$  are indexes.  $1 < i <$  Total number of agents Information Management.  $1 < j <$  Total number of Sensor Agents.  $0 < k <$  Total number of Actuator Agent.

**Fig. 2** Relationship between agents and information flow

The hardware of each IMA was designed in 3 sections (wireless communication using the ZigBee protocol processing stage and a power stage). Each stage is an electronic boards that fit together one after another to form the IMA, see Fig. 3.

**Fig. 3** Hardware of AGI information management



### 3 Learning Method Proposed

The data from the human-system interaction was obtained. The next step was recording these over a period of time and establish strategies to turn the data into useful information (each data was recording, see Fig. 4). In this way it is possible to find preferences of the user. It was decided to consider human behavior and his reaction to temperature changes as particular rules for each person.

We decided made a learning of rule user. In this case, we propose the following algorithms: Back-Propagation Neural Network, Minimum Distance Classifier, Naive Bayesian Classifier and Classifier K Near Neighbor (KNN).

It was registered the data during 14 days (at month September-October). We have 31,500 records of sensors and actuators. The data collected in the database were transformed to make them useful information for the classifier, ie, the feature vectors were obtained. We formed two databases feature vectors, one with 21 features and other features 9. The 21 characteristics that were taken were: month, day, hour, average temperatures 7 sensors in different parts of the room, the external average temperature, average humidity 7 sensors, the external average humidity and differences in temperature and humidity. The 9 characteristics of the other database includes the same features as above, but 7 temperature sensors are combined to form one temperature, in the same way the 7 humidities are combined. We use the state of the four actuators laboratory window to set each label.

The proposed learning method described below:

1. Obtaining feature vectors with labels
2. Determine number of vectors in the database, calculate number of classes and number of vectors per class.
3. Obtaining feature vectors with labels
4. Determine number of vectors in the database, calculate number of classes and number of vectors per class
5. Sort the classes by number of vectors (high to low), take classes by adding the number of vectors for each to get 70 % of the database.
6. Training classifier with frequent new database classes.
7. Stage proactivity. Environmental conditions are detected in the environment, these form a new input vector. This vector will be introduced in the classifier trained. The classifier will throw the classification label, which is the configuration status of the actuators. In this way the system must move the actuators in the real environment.

Date	Hour	Id Element	Data
------	------	------------	------

**Fig. 4** Elements of the data base

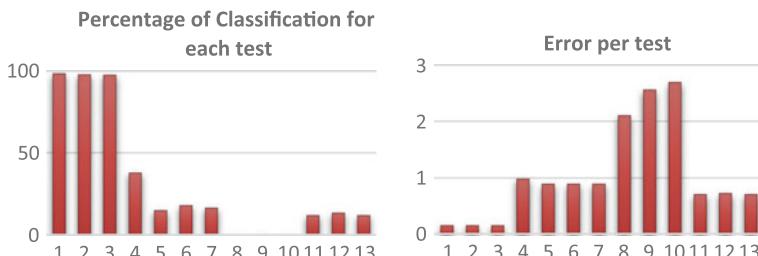
## 4 Experimental Tests and Results

The tests were performed for a single user during 14 days (September-October) in the room. The user manipulates the system using the graphical interface and the user reacts to changes in the environment. 13 different tests were performed using the Back Propagation neural network. The tests were conducted with different values of learning rate constant and different network structure. The cross validation method was used to compare the performance of neural networks. 10 folders were performed for comparison. It was performed 33 executions of the program for each folder because the weights and thresholds are initialized via a search engine pseudorandom. 100,000 iterations maximum and minimum error 0.2 were established.

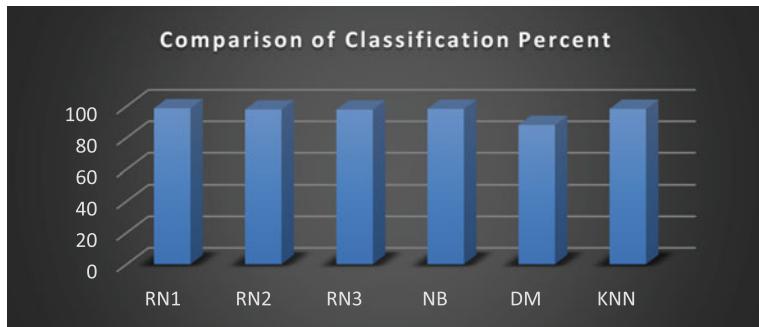
We use the database of 9 characteristics for the test 1–7. We use the database with 21 characteristics for test 8–13, see the results in Fig. 5. The performance of tests 1, 2 and 3 were higher than 90 %. In the test 1, the neural network was set with 9. In the test 1 neurons in the input layer, 17 neurons in the hidden layer and 1 neuron in the output layer. In the test 2, the neural network was set with 9 neurons in the input layer, 17 neurons in the hidden layer and 3 neurons in the output layer. In the test 3, the neural network was set with 9 neurons in the input layer, 17 neurons in the hidden layer and 12 neurons in the output layer. These three test was made with a constant speed adaptive of 0.2.

Three neural networks were selected because the three results had the best performance. We compared the three networks with the Naive Bayes classifier, the Minimum Distance classifier and the KNN classifier, see Fig. 6. The results shows classifications higher than 90 %, except the minimum distance that got a percentage of average classification of 83 %, it was the lowest of the six classifiers.

The proposed strategy was tested in order to measure its potential. For this purpose we decided to design test scenarios to consider different environmental conditions due to changes in the climate of Leon, Guanajuato, Mexico. We set different patterns of a reaction user about environment. 4 scenarios was proposed (winter season, warm atmosphere, rainy days and an ideal environment were chosen in the scenarios for a specific comfort). The results using a data base of 21



**Fig. 5** Results using neural networks with collected user data



**Fig. 6** Percentage classification with collected user data

characteristics were not satisfactory, for this reason we decided not to use it. In the next test we use the data base of 9 characteristics.

We run 13 test (with the same structure and learning rate constant that 13 test above) with the proposed scenarios. The yields of the 13 tests were deficient because ratings less than 20 % were obtained, the neural networks not converged with the proposed vectors. 6 tests were performed using Batch Propagation neural network with Momentum, but the percentages classification showed no improvement. The test (using scenarios) with the algorithms Naive Bayes, minimum distance and KNN had a classification yield of 98.66 %, 58.65 % and 88.65 % respectively.

## 5 Conclusion and Future Work

In this work, it is presented a system to help people with a physical disability to get comfort in any room of a house, this due to the actuators can be controlled by a mobile device. In the same time, the system detects thermal comfort preferences for give an automatic assistance.

In first results, the neural networks show classification higher percentages of 90 % (in the 1, 2 and 3 test). In the test, we can see that the neural networks have to adjust various parameters to try to achieve good ratings. In a process of continuous learning is possible to get new patterns and set parameters again. When increasing the number of classes for other environmental conditions, the results obtained were unfavorable with ratings less than 20 % (using proposed scenarios). Neural networks cannot be a good solution for learning user rules (considering the characteristics used). Therefore was selected Naive Bayes Classifier because had classifications more than 90 % for both user data monitored for 14 days or for the scenarios proposed.

As future work we propose to make a study using corporal temperature of the user. It is also hoped that other future research can be performed with data

collection of a year (or prolonged time) as learning research behavior of users, whose changes in habits can be influenced due the changes of the seasons, the climate, external information as personal agenda, holidays that celebrates the user and even consider implementing health sensors in the user. It is important to consider that can be performed usability studies to see if the system is accepted by users in real conditions.

**Acknowledgments** The authors want to acknowledge the kind and generous support from CONACyT and DGEST to this project.

## References

1. Aztiria, A., Augusto, J.C., Basagoiti, R., Izaguirre, A., Cook, D.J.: Learning frequent behaviors of the users. *IEEE Trans. Intell. Environ. Syst. Man Cybern. Syst.* **43**(6) (2013)
2. Ducatel, K., Bogdanowicz, M., Scapolo, F., Leijten, J., Burgelman, J.C.: Scenarios for ambient intelligence. IST Advisory Group. Technical Report Feb. 2001. (2010). <ftp://ftp.cordis.lu/pub/ist/docs/istagsenarios2010.pdf>
3. Vainio, A.M., Valtonen, M., Vanhala, J.: Proactive fuzzy control and adaptation methods for smart homes. *IEEE Intell. Syst.* **23**(2), 42–49 (2008)
4. Cook, D.J., Das, S.K.: How smart are our environments? An updated look at the state of the art. *Perv. Mobile Comput* **3**, 53–73 (2007). Elsevier Science
5. Heierman, E.O., Cook, D.J.: Improving home automation by discovering regularly occurring device usage patterns. In: Third IEEE International Conference on Data Mining, pp 537–540 (2002)
6. Rao SP, Cook DJ (2004): “Predicting inhabitant action using action and task models with application to smart homes”. *Int J Artif Intell Tools (Architectures, Languages, Algorithms)* **13** (1):81–99
7. Aztiria, A.; Izaguirre, A.; Augusto, J.C. (2010), “Learning patterns in ambient intelligence environments: a survey”, Springer Science+Business Media B. V., pp 35–51, 23 May 2010 doi:[10.1007/s10462-010-9160-3](https://doi.org/10.1007/s10462-010-9160-3)
8. Stankovski, V., Tmkoczy, J.: Application of decision trees to smart homes. In: Augusto, J.C., Nugent, C.D. (eds.) Designing Smart Homes. The role of artificial intelligence, m1: copyright 2006, The Institution of Engineering and Technology, pp 132–145. Springer (2006)
9. Youm, S., Lee, G., Park, S., Zhu, W.: Development of remote healthcare system for measuring and promoting healthy lifestyle. In: *Expert Systems with Applications*, vol. 38, pp. 2828–2834. Elsevier (2011)
10. Botia, J.A., Villa, A., Palma, J.: Ambient assisted living system for in-home monitoring of healthy independent elders. In: *Expert Systems with Applications*, vol. 39, pp. 8136–8148. Elsevier (2012)
11. Mowafey, S., Gardner, S., Patz, R.: Development of an ambient intelligent environment to facilitate the modelling of “Well-being”. In: IET Seminar on Assisted Living 2011, pp. 1–6 (2011)
12. Mowafey, S., Gardner, S.: Towards ambient intelligence in assisted living: the creation of an intelligent home care. In: Science and Information Conference 2013, London, U.K., pp. 51–50 (2013)
13. Suryadevara, N.K., Gaddam, A., Rayudu, R.K., Mukhopadhyay, S.C.: Wireless sensors network based safe home to care elderly people: behaviour detection. In: *Sensors and Actuators A* **186**, El Sevier, pp. 277–283 (2012)

14. Goribar, H.: Fundamentos de aire acondicionado y refrigeración. Limusa. Noriega Editores, Mexico (2009)
15. Sandra, L., Rosario, B., Ángel, C.M., Víctor, Z., Francisco, M.J., Arnulfo, A., Guillermo, M.: Physical implementation of an customisable system to assist a user with mobility problems. innovation in medicine and healthcare 2015, Volume 45 of the series Smart Innovation, Systems and Technologies pp 65–74. doi:[10.1007/978-3-319-23024-5\\_6](https://doi.org/10.1007/978-3-319-23024-5_6)
16. Mendez, G., Casillas, M.A., Baltazar, R., Ramírez, C.L., Mancilla, L., Lopez, S.: Intelligent management system for the conservation of energy. Intell. Environ. 120–123 (2015)
17. Barrón Llamas, T.J. Baltazar, R. Casillas Araiza, M.A., Cuesta Frau, D., Zamudio Rodríguez, V.M.: Comparison study of micro-algorithms for lighting comfort. In: International Conference of Numerical Analysis and Applied Mathematics 2014(ICNAAM 2014), Rhodes, Greece, 22–28 September 2014, AIP Conference Proceedings, vol. 1648, p. 820007 (2015). doi:[10.1063/1.4913026](https://doi.org/10.1063/1.4913026). ISBN 978073512873

# **Design of a Middleware and Optimization Algorithms for Light Comfort in an Intelligent Environment**

**Teresa Barrón Llamas, Rosario Baltazar, Miguel A. Casillas,  
Lenin Lemus, Arnulfo Alanis and Víctor Zamudio**

**Abstract** The evolution of technology allows to people with special capabilities of mobility to perform the activities faster and easier. The intelligent environments combined with optimization algorithms and middleware agents could help to this aim. This paper presents the design and the implementation of an architecture of a middleware agent that allows us to make the communication between heterogeneous devices (sensors and actuators of different communication protocols from WiFi to ZigBee). On the other hand, we present a comparison study between micro-algorithms used to get lighting comfort in order to perform an activity in a confined space; this is affect by the light from the outside, which can be blocked by shutters and doors, and lighting of lamps obtained within this space. The micro-algorithm evaluated were: Genetic Algorithm (GA), Artificial Immune System (AIS), Estimation Distribution Algorithm (EDA), Particle Swarm Optimization (PSO), Bee Algorithm (BA) and Bee Swarm Optimization (BSO).

**Keywords** Optimization micro-algorithms • Middleware agent • Intelligent environments • Light comfort

---

T.B. Llamas · R. Baltazar (✉) · M.A. Casillas · V. Zamudio

Instituto Tecnológico de León, León, Mexico

e-mail: rosario.baltazar.f@gmail.com; r.baltazar@ieee.org

T.B. Llamas

e-mail: barronllamas@gmail.com

M.A. Casillas

e-mail: miguel.casillas@gmail.com

V. Zamudio

e-mail: vic.zamudio@gmail.com

L. Lemus

Universitat Politècnica de Valencia, Valencia, Spain

A. Alanis

Instituto Tecnológico de Tijuana, Tijuana, Mexico

## 1 Introduction

The basic idea behind Ambient Intelligence (AmI) is that by enriching an environment with technology (mainly sensors and devices interconnected as agents), a system can be built to take decisions to benefit the users of that environment based on real-time information gathered and historical data accumulated [1]. With the creation of Intelligent Environments, people with mobility difficulties can be supported on their tasks or activities, because the environment where they meet provide comfort physical and/or omission of physical activities that AmI be able to perform.

Also in Čongradac raises as goal corroboration, by a simulation, increasing the energy efficiency of a system of adjustable lighting in a typical room, using a GA with obtaining direct light components [2]. A comparison of GA and BSO for obtaining lighting by lamps necessary for the completion of an activity [3]. We, based on this research, add lighting obtained from outside through windows and doors that are affected by time of day and state of blinds.

There have been several studies on the use of micro-algorithms has shown good performance in solving optimization problems outperforming algorithms that are not micro, as in the case of [4]. Another comparison made with micro-algorithms is proposed in [5], which shows the use of a micro-algorithm based on clonal selection theory for solving numerical optimization problems. Due to it is difficult to obtain the distribution of the intensity of light and the optical design of LED lamps, in [6], using the micro-genetic algorithm to determine the optimum angle of a planar prism and the best distribution of the intensity in these lamps.

This paper presents the design of a middleware that allows us to communicate with different agents (sensors and actuators) in a smart environment. In addition, the middleware developed helps us make management information database, which users make management; this information is sent from the sensors to the computer and the computer if you have to make modifications to the environment sends information to actuators. To determine if the environment should be modified we use optimization micro-algorithms.

## 2 Intelligent Environments to Support in Activities to People with Mobility Problems

Ambient Intelligence refers to smart Software and Hardware that helps people in their daily lives; the AmI may be related to areas such as Artificial Intelligence, Sensor and Actuator Networks, Computer Networks and Middleware, Ubiquitous Computing (Ubicomp) and Human-Computer Interaction (HCI) [7].

Operating an AmI system begins with obtaining environment information using sensors. This information is transmitted by a network and is pre-processed by what is called middleware, the middleware provides the bridge between the different technology sensors/actuators and the rest of system [8, 9]. To make decisions about

the changes in the environment, then a comparison algorithms, historical data and/or humans is necessary. Having made a decision, whether to modify the environment in a similar way we obtained the information from the sensors, information is sent to the actuators to modify the environment.

With the creation of an AmI, we obtain an atmosphere of comfort according to the activities carried out within a closed space. The use of AmI allows us to make proper use of available resources, that is, without wasting. In this project devices that help us manage the lighting in a room used to help people with mobility problems. To reach this objective we compared several algorithms to design an Intelligent Environment that helps to determine or give output parameters that will provide interaction of the ambient with the user, or simply change the ambient.

The optimization algorithms used in this research were: Genetic Algorithm (GA) is a technique for solving search and optimization problems based on the theory of evolution of species and natural selection [10]. Artificial Immune Systems (AIS) are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models, which are applied to solving problems [11].

Estimation of Distribution Algorithm (EDA) in contrast to GA's, in the EDA's don't have crossover and mutation operators. Instead, the new population of individuals is sampled from a probability distribution that is estimated from a database comprised of individuals from previous generations [12].

Particle Swarm Optimization (PSO) is tied to artificial life (A-life), swarming theory in general; it is also related with evolutionary computation, and has ties to both genetic algorithms and evolutionary programming [13]. Bees Algorithm (BA) is an optimization algorithm inspired by the natural foraging behavior of honey bees to find the optimal solution (nectar) [14]. Finally, the Bee Swarm Optimization algorithm (BSO) is a meta-heuristic hybrid between the PSO (discussed below) and BA algorithms. The main idea of BSO is based on taking the best of each meta-heuristic to get better results than those obtained any of them [15].

The aforementioned algorithms were programmed as micro-algorithms. The micro-algorithms have basically the same functionality as the standard algorithms; the only difference is that use small populations of less than 10 individuals [16]. The main objective of the use of micro-algorithms in this work is to find an optimal solution to the problem already raised and reduce processing time [17].

### 3 Agents Connected in a Multi-agent Architecture

To get information for Intelligent Environment, we use agents that measure environmental variables of light or another kind that contributed to the support to people with mobility problems. In this paper the detection, measurement and data collection of the environment is processed within of the architecture of the middleware agent where the information is coming true heterogeneous devices. One type of devices is a microcontroller NanoCore12 [18] that has a MC9S12CMCU another is the microcontroller LaunchPad MSP430 [19]. This plate work to 16 bits, it is

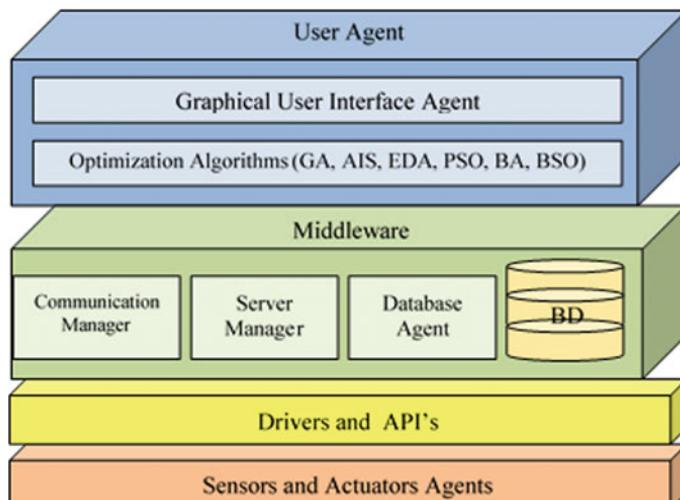
inexpensive and includes its own hardware interface for programming and debugging both of them have a light sensor, the middleware agent takes the value of the sensor and it is transmit to a destination, which may be another device. Another device is a Sun SPOT: (Sun Small Programmable Object Technology) is small wireless device developed by Sun Labs [20]. This device includes specific applications for sensors and actuators, such as accelerometers, light sensors, temperature sensors, LEDs, buttons and general pin I/O. The Mote IRIS has a Radio Processor Module IRIS or XM2110: (2400–2483.5 MHz) uses the Atmel RF230 is compatible with IEEE 802.15.4, has an integrated RF transceiver Zig-Bee with a microcontroller atmega1281. These motes have an integrated open source operating system based on embedded components called TinyOS [21].

Finally, the middleware has the capability to get the data from sensors, including cameras, microphone, accelerometer and touch screen of tablets and smartphones, in order to interact with the intelligent environment. These devices or agents have capacity to receive data from the sensors or other agents and send information to actuators.

All of these devices are connect to obtain measurements of lighting in an enclosed space; these measurements are use as input variables of the algorithms.

## 4 Middleware in Intelligent Environments

In this work a middleware that was designed to interact with different types of devices is presented. The devices have different operating systems and communications protocols. They are connected directly to the computer through a



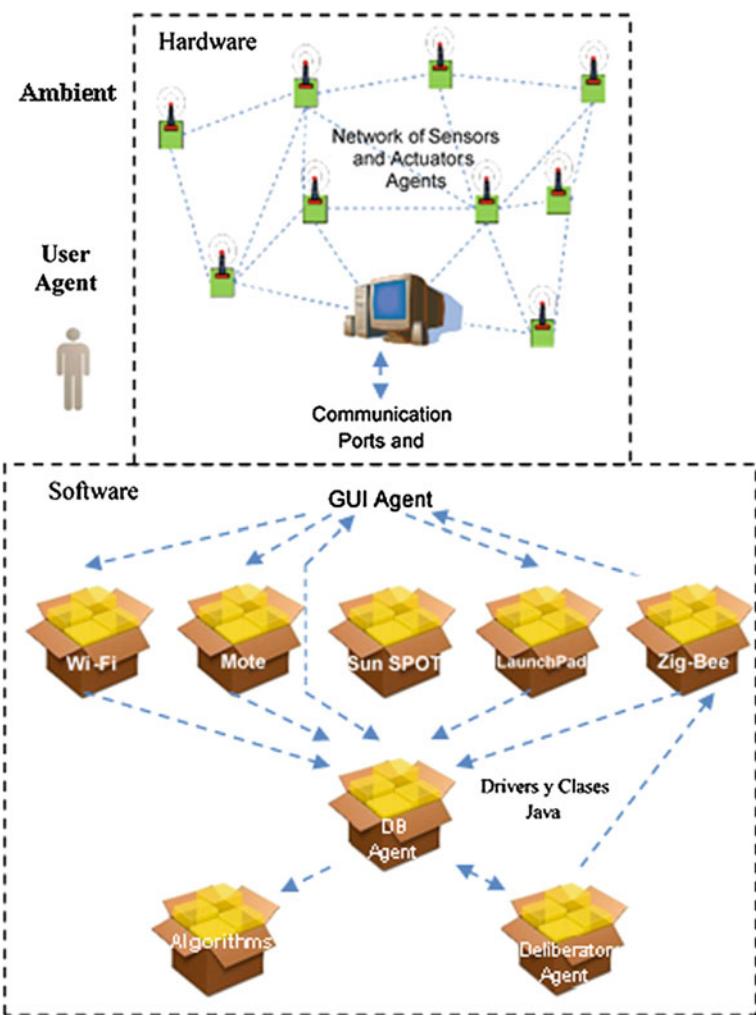
**Fig. 1** Middleware agent architecture

middleware designed to this purpose. Figure 1 shows the Middleware Architecture used. *Communication Manager* Module is responsible for send/receive information to/from users and different devices connected to the computer where the application and database are located, as well as to format the information. The job of the *Server Manager* module is responsible for starting and stopping servers needed to initiate communication with some devices. The information received/sent to users/devices must be stored for future reference or analysis, which the *Database Manager* module handles. This module is also responsible for providing information required by the optimization algorithms. Drivers and APIs are needed software that allows your computer to communicate with devices correctly; this software can be installed automatically if the computer recognizes the device, if not, can be found on the official websites of the devices or CD included if purchased as a KID. The sensors used for this jurisdiction lighting project that are embedded in different devices mentioned above.

In Fig. 2 illustrates how the AmI works when set to run the application (on a server or a normal computer) and devices are also turn on. The sensors begin to send data with their respective communication protocols and these give a format to the data according to how each device sends its information. Once you set a data format, they are sent to objects in the database that stores. Then, only certain information is sent to Algorithms objects by Database objects. After obtaining a solution of the conditions in which the ambient should be, deliberator agent consult information from database and determine if the ambient must be modified, if this happens, the deliberator agent send a MySQL instruction to the database to store instructions on modifying the ambient. It is ready to connect actuators that routed to the proper format with the instructions on modifying the environment. Some experiments to obtain the information of the devices were made to establish communication with the middleware and send the data to the Database. The user through a GUI can make the starting and ending of the communication. This interface has the images of the five different devices or agents to start the acquisition of the data.

## 5 Testbed and Algorithm Parameters

Testbed had a surface 12.5 m long and 8.7 m wide. It contains 21 lamps (L), 4 windows whit their blinds (P), 1 door (D), and 21 workspaces or zone (Z). The possible activities are read, computer work, stay or relax, project, and exposition, see Table 1, in this table shows the luminosity needed to perform an activity inside of the laboratory. The required illumination in each activity had a tolerance of  $\pm 50$  luxes.



**Fig. 2** AmI functionality. *Hardware* network includes sensors/actuators and the equipment or server where the devices are connected. *Software* represents the organization of Java classes and how they communicate between them with the GUI made with the framework ZK

**Table 1** Lighting recommended for each activity [4]

Activity	Description	Average lighting service (lux)
1	Read	750
2	Computer work	500
3	Stay or relax	300
4	Project	450
5	Exposition	400

## 5.1 Objective Function

Getting comfort lighting is need a mathematical representation of the test scenery, which we obtain only the luminosity needed in a specific moment of time and according to the users inside. The size of the individual (TI) used to execute in each micro-algorithm is given by Eq. (1) and the visual representation is given in Fig. 3.

$$TI = (P * B_p) + (D * B_d) + (L * B_l) \quad (1)$$

where:

- $P, D$  and  $L$ : is the number of blinds, doors and lamps (respectively)
- $B_p, B_d$  and  $B_l$ : is the number of bits representative of each blind, door and lamp in the individual or particle. The decimal value of the binary representation blind, door and lamp will help us determine the percentage of light that is coming from the outside to a specific area of the laboratory.

The aim is to obtain the required lighting for users working in specific zones of the Testbed this to maximize the use of natural light. The objective function to determine the required lighting in the laboratory is given by Eq. (2) and the objective function for each zone is given by Eq. (3).

$$F = \sum_{z=0}^{Z-1} abs(F_z) \quad (2)$$

$$F_z = I_{az} - I_{pz} - I_{dz} - I_{lz} \quad (3)$$

where:

- $F$ : is the total illumination required in the laboratory.
- $F_z$ : is the total illumination that a user will receive given by combination of luminous flux of the different light sources and the activity being performed in the zone.
- $I_{az}$ : is the luminous flux that is required to perform an activity specified by the user in a zone.
- $I_{pz}, I_{dz}$  and  $I_{lz}$ : is the luminous flux that blinds, doors and lamps (respectively) allow moving into a zone of the laboratory, It is given by Eqs. (3), (4) and (5) respectively.

$P * B_p$	$D * B_d$	$L * B_l$
0   0   ...   0   0   0   0   ...   0   0   0   0   0   ...   0   0   0		

**Fig. 3** Size of the individual in the algorithm

To determine the objective function are:

$$I_p = \sum_{p=0}^{P-1} \frac{F_{pzv_p} * S_p}{100} \quad (4)$$

$$I_d = \sum_{d=0}^{D-1} \frac{F_{dzv_d} * S_d}{100} \quad (5)$$

$$I_l = \sum_{l=0}^{L-1} \frac{F_{lz} * v_l}{V_l} \quad (6)$$

where:

- $F_{pzv_p}$  and  $F_{dzv_d}$ : are the flow rate respectively of a blind  $p$  or a door  $d$  allow moving into to a zone  $z$  according to the decimal value  $v_p$  and  $v_d$  respectively.  $v_p$  and  $v_d$  are obtained from the binary representation of a blind  $B_p$  and  $B_d$  of an individual/particle respectively.
- $S_p$  and  $S_d$ : is the luminous flux that a sensor located near of a blind  $p$  and door  $d$  is measured.
- $F_{lz}$ : is the luminous flux that a lamp  $l$  will provide a zone  $z$ .
- $v_l$ : is decimal value obtained from the binary representation of a lamp  $B_l$  of an individual.
- $V_l$ : is the maximum decimal value from the binary representation of a lamp  $B_l$  of an individual.

## 5.2 Results

Initialization parameters of the micro-algorithms were obtained experimentally. The distribution is as follows: Population size 5 individuals/particles, calls to function 2000, Elitism for GA, AIS was 20 % of the population. To create the Probabilistic Model of EDA was used 40 % of the population, speed for the BSO and PSO was 20, to search radius of the BA and BSO 5 individuals/particles were used and search for each individual/particle 5.

The real test scenario used in this comparison of micro-algorithms is as follows: “Tuesdays and Thursdays between 17:00 and 19:00,  $z_0$  and  $z_1$  are computer work zones; classes (exhibitions) covering the zones  $z_9, z_{10}, z_{11}, z_{12}, z_{13}, z_{14}, z_{15}, z_{16}, z_{17}$  and  $z_{19}$  are given. In total 5000 lux (500, 500, 400, 400, 400, 400, 400, 400, 400, 400, and 400) are required”.

The fitness ( $F$ ) obtained using the aforementioned parameters are listed in Table 2. To determine which micro-algorithm is the best, we use the Friedman test [22] using the SPSS program. The results of the application of this statistical test are show in Table 3.

**Table 2** Medians of the different tests

Iteration	Micro-algorithm					
	GA	AIS	EDA	PSO	BA	BSO
1	553.5625	549.848214	549.848214	327.839286	457.625	314.714286
2	536.303571	512.125	512.125	321.491071	421.910714	339.517857
3	586	586	586	302.991071	467.098214	305.848214
4	592.839286	585.339286	585.339286	346.642857	441.303571	306.446429
5	570.071429	570.071429	570.071429	332.848214	424.9375	312.017857
6	580.125	580.125	580.125	340.625	469.232143	318
7	513.366071	511.223214	511.223214	350.678571	424.651786	313.303571
8	587.723214	570.946429	570.946429	331.392857	404.616071	309
9	609.758929	609.758929	609.758929	293.526786	489.044643	306.928571
10	565.901786	559.044643	559.044643	292.625	441.25	310.089286

While it is known that our objective function (Eq. 2), we expect to have a difference of 0, we have considered a light range in each activity, in our case it is  $\pm 50$  luxes. Since there are 12 activities in a specific time range the values obtained by  $F$ , given the absolute, is from 0 to 600 luxes. Similarly, the values obtained in  $F_z$  (Eq. 3) are both positive and negative, so that when you set the status of the different objects will be a tradeoff between those in which the lighting is exceeded and those they lack.

**Table 3** Friedman test with asymptote

Statistics				Ranges			
Micro-algorithm	N. data	Mean	Deviation	Minimum	Maximum	Micro-algorithm	Average range
GA	10	569.5632	28.68020	513.37	609.76	GA	5.60
AIS	10	563.4482	31.75700	511.22	609.76	AIS	4.70
EDA	10	563.4482	31.75700	511.22	609.76	EDA	4.70
PSO	10	324.0661	21.10596	292.63	350.68	PSO	1.60
BA	10	444.1670	26.19620	404.62	489.04	BA	3.0
BSO	10	313.5866	9.91174	305.85	339.52	BSO	1.40

## 6 Conclusions

This work allows us to obtain a state of comfort regarding lighting in various activities within the Testbed. To make the integration of different types of devices was necessary to know how they worked each and how to communicate with the team for the storage of information. This was possible with the implementation of a middleware that allowed us binding diversity of components (software and hardware).

In addition, micro-algorithms have been proven to be useful approach in order to find optimal solutions in short periods to various types of problems. Because not all algorithms are useful for all kinds of problems it is necessary to make comparisons between them. We developed an objective function that helps us to get ambient comfort according to information sent from devices and execution of algorithms. Using the Friedman test obtained that the samples of each data set are different and the micro-algorithm that gives us a better combination of the state in which blinds, door and lamps have to be to get the appropriate lighting according to the activity performed was the  $\mu$ BSO. This is taking an asymptote of 0.05. We conclude that the BSO was the best micro-algorithm due to the combination of the method of updating the particles taken from PSO (which took second place), and the radius of search implemented in the best particles taken from BA.

As future work, we pretends to include other types of sensors (humidity, pressure, temperature, etc.) and actuators (door locks and windows) programs, which are linked by the middleware.

Include context for understanding the environment is able to automatically detect the activities that are being carried at a particular time and place or area.

This work can be implemented in other enclosed spaces such as homes, offices, public facilities, etc. This implementation helps to save energy and improve lighting comfort. Moreover, thanks to its modularity, it can improve by using new technologies emerging.

**Acknowledgments** We would like to thank CONACYT, ITL and DGEST for supporting this project.

## References

1. Augusto, J.C.: Ambient intelligence: basic concepts and applications. *Commun. Comput. Inf. Sci.* **10**, 16–26 (2008)
2. Čongradac, Velimir D., Milosavljević, Boško B., Veličković, Jovan M., Prebiračević, Bogdan V.: Control of the lighting system using a genetic algorithm. *Thermal Sci.* **16**, 237–250 (2012). doi:[10.2298/TSCI120203075C](https://doi.org/10.2298/TSCI120203075C)
3. Romero, W., Zamudio, V., Baltazar, R., Sotelo, M., Soria, J.: Comparative study of BSO and GA for the optimizing energy in ambiente intelligence. In: *Advances in Soft Computing, MICAI 2011, Part II, LNAI 7095*, pp. 177–188 (2011)

4. Romero, L.A., Zamudio, V., Sotelo, M., Baltazar, R., Mezura, E.: A comparison between meta-heuristics as strategies for minimizing cyclic instability in ambient intelligence. *Sensors* (2012)
5. Herrera-Lozada, J.C., Calvo, H., Taud, H.: A micro artificial immune system. *Polibits* **43**, 107–111 (2011)
6. Kim, Y.S., Choi, A.S., Jeong, J.W.: Applying microgenetic algorithm to numerical model for luminous intensity distribution of planar prism LED luminaire. *Opt. Commun.* **293**, 22–30 (2013)
7. Augusto, J.C., Callaghan, V., Cook, D., Kameas, A., Satoh, I.: Intelligent environments: a manifesto. *Hum. Centric Comput. Inf. Sci.* **3**, 12 (2013). Springer
8. Baquero, R., Rodríguez, J., Mendoza, S., Decouchant, D.: Towards a modular scheme for the integration of ambient intelligence systems. In: 5th International Symposium on Ubiquitous Computing and Ambient Intelligence UCAMI (2011)
9. Youngblood, G.M.: Middleware, Smart Environments: Technologies, Protocols, and Applications. In: Cook, D.J., Hoboken, S.K. (eds.), pp. 101–127. Wiley, New Jersey (2005)
10. Araujo, L., Cervigón, C.: Evolutionary Algorithms, A Practical Approach, 1st edn. In: de C.V. (ed.). SA Alfaomega Grupo, Mexico (2009)
11. Leandro, N.: de Castro y Jonathan Timmis. In: Artificial Immune System: A New Computational Intelligence Approach, 1st edn. Springer, UK (2002)
12. Larrañaga, P., Lozano, J.A.: Estimation of distribution algorithms (2003). <http://www.redheur.org/sites/default/files/metodos/EDA01.pdf>
13. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, Perth, Australia. IEEE Service Center, Piscataway, NJ, (1995) (in press)
14. Pham, D.T., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., Zaidi, M.: The bee algorithm—a novel tool for complex optimisation problems. *Intell. Prod. Mach. Syst.* **2** (2006)
15. Sotelo-Figueroa, M.A., Baltazar-Flores, M.D.R., Carpio, J.M., Zamudio, V.: A comparison between bee swarm optimization and greedy algorithm for the knapsack problem with bee reallocation. In: 2010 Ninth Mexican International Conference on Artificial Intelligence (MICAI), pp. 22–27, 8–13 Nov 2010
16. Tam, V., Cheng, K.Y., Lui, K.S.: Using micro-genetic algorithms to improve localization in wireless sensor networks. *J. Commun.* **1**(4) (2006). Academy Publisher
17. Beaubrun, Ronald, Ruiz, Jhon Fredy Llano, Poirier, Benoit, Quintero, Alejandro: A middleware architecture for disseminating delay-constrained information in wireless sensor networks. *J. Netw. Comput. Appl.* **35**, 403–411 (2012). doi:[10.1016/j.jnca.2011.09.002](https://doi.org/10.1016/j.jnca.2011.09.002)
18. Technological Arts, Fundada in 1995 (2014). <http://www.nanocore12.com/>
19. Oracle (2013). <http://www.sunspotworld.com/Tutorial/index.html>
20. MEMSIC Inc (2013). [http://www.memsic.com/userfiles/files/datasheets/wsn/iris\\_datasheet.pdf](http://www.memsic.com/userfiles/files/datasheets/wsn/iris_datasheet.pdf)
21. Texas Instruments (2013). <http://www.ti.com/tool/msp-exp430g2>
22. Derrac, J., García, S., Molina, D., Francisco, F.: A practical tutorial on the use of nonparametric statistical test as a methodology for comparing evolutionary and swarm intelligence algorithms. In: Swarm and Evolutionary Computation. Elsevier (2011)

# **Autism Disorder Neurological Treatment Support Through the Use of Information Technology**

**Esperanza Manrique Rojas, Margarita Ramírez Ramírez,  
Hilda Beatriz Ramírez Moreno, Maricela Sevilla Caro,  
Arnulfo Alanís Garza and José Sergio Magdaleno Palencia**

**Abstract** Autism is a complex neurological disorder that typically lasts a lifetime. It is part of a group of disorders known as autism spectrum disorders. PPPs for Autism EdNinja, were developed by experienced therapists in the neurological disorder autism. This tool support in the treatment and development of skills a child with this disorder between the ages of 4–10 years. The company MindHUB developer of EdNinja APS is a company in the city of Tijuana, Baja California, Mexico; They APS these are developed considering the cultural characteristics of Mexico, not rear confusion in the child. To market to other countries this tool, you must do the translation and adaptation to the corresponding culture.

**Keywords** Autism • Information technology • APPS

---

E.M. Rojas (✉) · M.R. Ramirez · H.B.R. Moreno · M.S. Caro  
Facultad de Contaduría y Administración, UABC, Calzada Universidad 14418,  
Parque Industrial Internacional Tijuana, C.P. 22390 Tijuana, BC, Mexico  
e-mail: emanrique@uabc.edu.mx

M.R. Ramirez  
e-mail: maguiram@uabc.edu.mx

H.B.R. Moreno  
e-mail: ramirezmb@uabc.edu.mx

M.S. Caro  
e-mail: mary\_sevilla@uabc.edu.mx

A.A. Garza · J.S.M. Palencia  
Departamento de Sistemas y Computación, Instituto Tecnológico de Tijuana,  
Calzada del Tecnológico S/N, Tomas Aquino, 22414 Tijuana, BC, Mexico  
e-mail: alanis@tectijuana.edu.mx

J.S.M. Palencia  
e-mail: jmagdaleno@tectijuana.edu.mx

## 1 Introduction

Autism is a complex neurological disorder that typically lasts a lifetime. It is part of a group of disorders known as autism spectrum disorders (ASD) in English. Currently diagnosed with autism is 1 in 68 people and 1 in 42 children, so it is the most common cancer, diabetes and pediatric AIDS combined. It occurs in all racial, ethnic and social groups and is four times more common in boys than in girls. Autism affects a person's ability to communicate and interact with others. It is also associated with routines and repetitive, such as obsessively arranging objects or following very specific routines behaviors. Symptoms can range from mild to very severe, Rodriguez [7].

If the child is diagnosed with autism, early intervention is essential for you to benefit the most of all existing therapies. Although parents may be difficult to label a child as "autistic", the former to faster diagnosis may act is done. Currently there are no effective means to prevent autism, no fully effective treatments or cures for this disorder. However, research shows that early intervention in an appropriate educational setting for at least 2 years during the preschool years can have significant improvements for many young children with autism spectrum disorders. As soon as autism is diagnosed, early intervention should start with programs focused on effective communication skills, socialization and cognitive development, Autism [2].

The MindHUB company Tijuana, Baja California, Mexico, has developed applications designed to help children in special education as autism and dyslexia among others these applications are integrated into a package called EDNINJA, which concentrates applications such as: sequences for autism, patterns, find the letters, sortings for autism, expressions for autism, these applications developed by experienced therapists in autism, specifically for children with special learning needs, EdNinja [4].

## 2 Autism Disorder Neurological

Tortosa [8], one of the main ways to help children with autism communicate better with others is the use of adapted systems. Currently, there is no standard treatment for autism spectrum disorder (ASD). There are several ways to maximize the child's ability to grow and learn new skills. The sooner you start, the more likely to have more positive effects on symptoms and skills. Treatments include behavior and communication therapies, skills development and/or medications to control symptoms.

Most parents of children with ASD suspect that something is not right when the child is 18 months old and seek help from 2 years old. Children with ASD are characterized by problems with:

- Simulation games
- Social Interactions
- The verbal and nonverbal communication

Some children seem “normal” before the first or second year of life and then have a sudden regression and lose language or social skills they had previously acquired. These symptoms can range from mild to severe, Amin [1]. A person with autism can be:

- Extremely sensitive about sight, hearing, touch, smell or taste (for example, you can refuse to wear clothes “that gives itching” and distress if forced to use it).
- Have unusual distress when routines are changed.
- Perform repeated body movements.
- Show an exaggerated attachment to objects.

Communication problems may include:

- You are unable to initiate or sustain a social conversation.
- Communicates with gestures instead of words.
- Develops language develops slowly or not at all.
- Do not adjust gaze to look at objects that others are watching.
- Do not refer to yourself properly (for example, says: “You want water” when they mean “I want water”).
- Do not point to direct the attention of others to objects (occurs in the first 14 months of life).
- Repeats words or memorized passages as commercial.

In the social interaction is another aspect that can be detected in children with ASD, clearly it shows the lack or difficulty making friends, play interactive games, retracts, it is possible to respond to eye contact or smiles or can avoid contact visual, you can treat others as objects rather spend time alone and not with others, it shows a lack of empathy. Early supports early detection strategies including verbal and nonverbal communication, contingency, taking turns, imitation and joint attention in children with autism. These strategies are based on five general principles.

Response to sensory information:

- Do not startle to loud noises.
- Has heightened or low senses of sight, hearing, touch, smell or taste.
- Normal noises that may seem painful and hold hands to his ears.
- You can avoid physical contact because it is very stimulating or overwhelming.
- rubbing surfaces, objects leads to the mouth or licking.
- Seems to have an increase or decrease in response to pain.

Game:

- Do not imitate the actions of others.
- Prefers solitary or ritualistic play.
- Displays or acted imaginative play.

Behaviors:

- It acts with attacks of intense anger.
- It is dedicated to a single topic or task.
- have a short attention span.
- has very narrow interests.
- Hyperactive or too passive.
- It shows aggression towards others or himself.
- Shows great need for equality.
- Uses repetitive body movements.

### 3 Technology Support Autism Treatment

The “Adaptive Technology” is used to define the field of action of technological care for people with disabilities. Rehabilitation Technology, Assistant Technology, Technology Access, or Assistive technology are some of them, Dietrich and Romano [3].

The intended for people whose disabilities cannot use the linguistic code oral-verbal communication systems it is called special communicators.

Adaptive technology can reduce the impact of disability and guarantee the right quality of life of people with special needs use of new technologies in autistic children, PRNoticias [6].

As Perez de la Maza [5], the most widespread use of ICT in the intervention with people with ASD says it focuses on five distinct areas:

1. Education
2. Communication
3. Leisure
4. Classification
5. Diagnosis

The company MindHUB Tijuana, Baja California, Mexico, has developed applications designed to help children in special education as autism and dyslexia among others. These applications are integrated into a package called EDNINJA, which concentrates applications such as: SEQUENCES FOR AUTISM, which teach everyday activities and skills by ordering sequences of pictures; PATTERNS Improve math skills by Identifying shapes, forms and colors; FIND THE LETTERS, that Improve reading skills by Identifying and coloring letters; SORTHINGS FOR AUTISM, Learn to Identify and sort Real-life objects in Their correct places or semantic fields; EXPRESSIONS FOR AUTISM, Train and Improve human emotion recognition by recreating facial expressions. These applications developed by experienced therapists in autism, specifically for children

with special learning needs. The main difference between applications for children with autism and other games is the sequence of detailed activities, as each is separated systematically, with daily and repetitive activities, making this ideal for therapists, special education teachers and support tool, EdNinja [4].

The characteristics of these applications shows the player's progress (child) to track per user, time and number of errors, you can choose between a male or female voice for instructions and tutorials, and is a bilingual application (English and Spanish).

Some advantages of this application:

- Improved communication with others.
- Improved temporal-spatial location.
- Improved ability to structure ideas and thoughts.
- Focus on self-help skills.
- Users can simplify daily activities in steps, improving their social skills.

EdNinja based applications development under the supervision of specialized therapists, this strongly supports the use of the tool as support treatment, and there is also another important element to consider: the cultural features of the country where this tool is used. For the company MindHUB, these APPS can market to other countries, you must make the acculturation and adaptation to the target country.

## 4 Conclusions

Increasingly, society takes greater awareness of the autistic problems. There is more integration for people with autism understand their scope is more and more tools to help them understand the world around them and communicate with others more effectively created.

Communication is one of the main factors that isolate children with autism in their environment. One of the great challenges of the families of those affected by this disorder and the professionals who adhere, once the diagnosis is known, is to influence their communication skills and in their personal and social autonomy. In this sense, new technologies are a very important platform and excellent support tool, and if used properly, can amplify the benefits that communication poses to the child's development.

“Tell me and I forget, show me and I remember, involve me and I learn.”  
Benjamin Franklin.

## References

1. Amin: Neurological Syndromes (2016). <http://sindromed.com/sindrome/sindromes-neurologicos/page/2/>. Accessed 26 Feb 2016
2. Autism speaks what is autism? A description (2016). <https://www.autismspeaks.org/qu%C3%A9-es-el-autismo>. Accessed 2 Feb 2016
3. Dietrich, L.S., Romano, S.L.: Adaptive Technology (2010). <https://sites.google.com/site/tecnologiasadaptativas2010/>. Consulted 5 Feb 2016
4. EdNinja (2015). <http://EdNinja.com/>. Accessed 4 Feb 2016
5. Perez de la Maza, L.: Computer applications for people with autism spectrum disorders. I Regional congress special educational needs: current and future challenges, pp. 391–397, Badajoz (2003)
6. PRNoticias (2015). <http://prnoticias.com/salud/20140046-las-tic-un-puente-de-comunicacion-para-los-ninos-con-autismo>. Cosnultado 9 Feb 2016
7. Rodriguez, M. (2015). <http://www.fchautismo.org/gaceta/16-que-es-el-autismo>. Accessed 16 Feb 2016
8. Tecno Autism (2015). <http://autismoytecnologia.webnode.es/investigando-/marco-teorico-autismo-y-nuevas-tecnologias/>. Accessed 10 Feb 2016
9. Tortosa, N.: Assistive technologies for people with autism spectrum disorders: a guide for teachers (2004). <http://diversidad.murciaeduca.es/tecnoneet/docs/autismo.pdf>. Accessed 10 Feb 2016

# **Information Technologies in the Area of Health and Use of Mobile Technologies in the Area of Health in Tijuana, Baja California**

**Margarita Ramírez Ramírez, Esperanza Manrique Rojas, Nora del Carmen Osuna Millán, María del Consuelo Salgado Soto, José Sergio Magdaleno Palencia and Bogart Yail Márquez Lobato**

**Abstract** Health is an issue of significant importance to any country, in Mexico, government institutions actions promoting advancement plans and implementation of strategies that allow the inclusion of information and communications technology in supporting been made diagnosis, treatments and procedures in public health. The computerization of the health sector ensures efficient processes and access to accurate, timely information at the right time, elementary factors in decision-making. The use of media such as cloud computing, storage and efficient access to patient information. What is the real situation observed in the city of Tijuana, Baja California. This study presents the opinion of health experts in the public sector, on the use and interest to use technological tools as support in substantial activities in the provision of health care.

**Keywords** Information technologies • Mobile technologies • Health

---

M.R. Ramírez (✉) · E.M. Rojas · N. del Carmen Osuna Millán  
Facultad de Contaduría y Administración, UABC,  
Calzada Universidad 14418, Parque Industrial Internacional,  
C.P. 22390 Tijuana, BC, Mexico  
e-mail: maguiram@uabc.edu.mx

E.M. Rojas  
e-mail: emanrique@uabc.edu.mx

N. del Carmen Osuna Millán  
e-mail: nora.osuna@uabc.edu.mx

M. del Consuelo Salgado Soto · J.S.M. Palencia · B.Y.M. Lobato  
Departamento de Sistemas y Computación, Instituto Tecnológico de Tijuana,  
Calzada del Tecnológico S/N, Tomas Aquino, 22414 Tijuana, BC, Mexico  
e-mail: csalgado@uabc.edu.mx

J.S.M. Palencia  
e-mail: magdalenopjs2@yahoo.com.mx

B.Y.M. Lobato  
e-mail: bogart@tectijuanc.mx

## 1 Introduction

Now speak of development in a society, it is to speak of progress in different areas of science and technology, and how these advances affect daily life, and enable humans to achieve better standards and quality of life.

Advances in science and technology change the lifestyle of human beings today most human activities revolve around the use of electronics and computing. You have electronic communication devices, it is very common, which offer applications and tools in daily life. These mobile devices have revolutionized the different communities, Smartphones (or smartphones) and tablets (or electronic tablets) have changed the way people consume information and communicate.

Access to information is made through mobile data connections, you may be connected to the Internet anytime, anywhere.

Health is one of the areas most should benefit from this revolution, every day new applications emerge whose objectives are to improve human well-being, increase information has aspects related to health and/or improve management and the control of diseases and health areas.

There are medical applications, which are software tools designed for use on mobile devices and offer a variety of services and utilities.

According to studies, it was estimated that by 2015, the sector of mobile applications related to health have been more than 500 million users [1].

The use of cloud computing is another area of opportunity to store and access data from anywhere and at any time, so it is possible to become an important support in the area of health.

This article presents an analysis of the use of information technologies in the area of health is presented, the current situation is analyzed through a descriptive study of doctors from the Mexican Social Security Institute, in the city of Tijuana, BC, and integration of cloud computing to support the storage of electronic medical records, sending information by this.

## 2 Research Background

Health is a topic of great relevance and a factor that determines and/or is determined by different factors such as the demographic situation, levels of health and education, hospital infrastructure; access to medicines, basic health services, among others.

You can identify breakthroughs in the health area; The National Development Plan 2013–2018 in Mexico, places health as one of its cross-cutting strategies. There are significant achievements such as increased life expectancy and decreased infant mortality in the country; However, progress is pending in some other key indicators in health development, as is the use of technology, which can support and promote progress and achievements in this area. In particular the development and

application of mobile technologies in the area of health can become a tool of a relevant value. The health promotion can not be without linking individuals with technology.

A tool of great importance that you can use in the area of health is cloud computing, which is identified by the Computer Society, as a paradigm in which information is stored permanently on Internet servers and sends the temporary client.

It is important to know the current situation regarding the issue of use and implementation possibilities of mobile technologies in the area of health as well as the diagnosis of needs, in particular the information was collected in a clinic in the Mexican Social Security Institute, in particular the clinic no. 36 of the Mesa de Otay, Tijuana, Baja California.

In 2009 President Obama used the American Recovery and Reconstruction Act (ARRA, the “Stimulus”) to set aside up to \$39 billion to fund adoption of electronic health records and to create health information exchanges along with a number of related programs to spur adoption and to pave the way for further development of the field. I feel it is also spawning a new wave of innovative entrepreneurial activity, most of it based in the cloud, the inspiration for the title of this book. Cloud computing has the goal of making computer resources available as they are needed some-what like other utilities we all rely on such as electricity and water. When combined with the increasing utilization of wireless and mobile technologies it offers a truly transformative platform for healthcare delivery [2].

### 3 Mobile Health Applications

Evaluate the results obtained in the teaching-learning of computer use, through the Clearly the smart phone has become a vital tool for humans, as well as connect via courier and social networks, we offer other services such as health care. Mobility is certainly support the work of health professionals in areas such as patient.

The smart phone has become a vital tool for humans, as well as connect via courier and social networks, we offer other services such as health care.

Mobile devices have acquired the status of a basic necessity rather than simply be an auxiliary device and luxury. Mobile networks are the platform more widespread communication that exists today and could be used as one of the most effective means for interconnecting providers, experts and other stakeholders to a large substantial number of consumers in the health system.

Mobility is certainly support the work of health professionals in areas such as patient tracking, management appointments and personalized management of medical records and electronic prescriptions. You can in turn become a practical method as hospitals strategies, in which the roles are avoided and allow storing medical records. Some organizations like the FDA (Food and Drugs Administration) US warn of the need to regulate this market applications addressing the

possible impact that the use of them may have on the health of citizens. The FDA has made a classification of mobile health applications into two broad:

1. Applications for consultation and storage of information generated by devices and clinical information systems.
2. Applications on a Smartphone, to make this work as a medical device, which can incorporate sensors such as an electronic stethoscope or show radiological diagnostic images.

Moreover Deloitte Center for Health Solutions, conducted a study based on the new trends of a change in the habits of consumption of health services.

Among the findings highlighted that:

- Latin America is one of the world's highest per capita number of mobile telephony.
- Applications that allow you to monitor your heart rate, measuring the amount of water ingested in one day, remember taking medication, my health offer users more autonomy and control of their welfare, and are already part of the routine of many people.

In Mexico, 107 million people have access to mobile phones, of which 50 % has a “Smartphone”, and uses applications [3].

Every day is more common term of mHealth, which describes the practice of medicine and public health through mobile devices in Latin American countries.

There is evidence that in Mexico, health applications rank fourth in shock, leaders and experts in the field of health have mentioned that technological applications in health are indispensable to increase addiction among users instrument; why it is important, that working intensely on developing mobile tools to prevent and treat various diseases [4].

## 4 Cloud Computing

They have made efforts to assist in Baja California the development of Science and Technology, among the achievements of these efforts we can identify the development cluster, among which are: the cluster of information technology and communication Cloud Computing. It is the term used to describe the possibility of using computer programs and resources from the internet. This is one of the main trends of development of computer applications as it increases the number of services offered, generating benefits and advantages such as the availability of information immediately.

Cloud computing is a popular model to provide technology-related services. Several enabling the Access on demand, through the net, to a set of shared resources [5].

## 5 Instituto Mexicano del Seguro Social (IMSS)

The IMSS is the institution with the largest presence in health care and social protection of Mexicans since its founding in 1943, for it combines research and medical practice, management of resources for the withdrawal of its policyholders. Today, more than half of the Mexican population, has something to do with school, so far, the largest of its kind in Latin America [6].

## 6 Methodology

This study was conducted through desk research and exploratory and descriptive study, which presents the opinion of a group of practicing general practitioners who care for patients in a clinic in the Mexican Social Security Institute.

Application technique directed questionnaire was used, the design of the questionnaire was using the Likert scale and Cronbach's alpha technique was applied to the validation [7].

## 7 Results

In descriptive exploratory research it was conducted in IMSS Clinic No. 36, located in Otay Mesa. Tijuana, located in the Technological Avenue Unnumbered, Otay, Tijuana, Baja California.

This clinic has eleven offices for general medical patients, in which patients are treated in both, serving approximately thirty patients daily for office.

A questionnaire to thirty doctors, who are attached to this clinic was applied, the most significant results obtained in this instrument are:

Thirty doctors who apply the questionnaire, 19 of them were males and 11 females. 87 % have mobile and there is still 13 % that does not use mobile phones.

47 % know of any mobile application in the area of health and 53 % do not know any mobile application in the health area.

The question was asked if they would like appointments to the consultations were managed by an application, to which they replied:

13 % that never interested them, 30 % might be interested, 34 % definitely interested and 23 % would not be interested.

To the question you like clinics incorporate mobile technologies to track the patient care?

56 % believe that if interested in this type of application.

I think 17 % is not necessary, 20 % think it might be good, 7 % think not interested.

Asked Considers that better control in the clinic would, if incorporated technological applications?

40 % think that if would take better control implemented this type of application.

47 % think that perhaps a better control in the clinic would be achieved.

13 % think that not necessary this type of application.

When asked technologies has made applications to diagnose diseases?

3 % answered yes using any application.

97 % answered not using any application.

To the question would use mobile applications for patient information?

57 % mention whether be interested in using mobile applications for patient information.

20 % think that maybe used.

23 % think that not interested to receive information by these means.

## 8 Conclusions

In general, mobile applications do not involve a significant advance from the point of view of medical technology. The most important achievement is to get facilitate and standardize the use of devices and applications by more natural interfaces.

Achieving highlight the potential of mobile health in improving disclosure of facilities based on health, such as maintenance schedules and appointments, send and receive information on patient treatment and diagnostic.

Using cloud computing to provide greater mobility and ease of access to real-time patient data from anywhere using any device.

We conclude that a smartphone in the pocket of her robe a doctor becomes a tool that will allow you to access and manage clinical patient information at any time and from any location.

It is important to adopt these technologies to advance and meet the standards and quality levels required in the health of our population tools.

## References

1. Bellocchio, S.: Everis Health, Salud 2.0: La innovación se toma la medicina, [www.everis.com/chile/WCLibraryRepository/.../everis%20Health.pdf](http://www.everis.com/chile/WCLibraryRepository/.../everis%20Health.pdf) (2011)
2. Braunstein, M.L.: Health Informatics in the Cloud
3. Deloitte, Encuesta Global 2011 de Consumidores de la Atención de la Salud Principales Hallazgos
4. Ramírez, R.: Tecnologías móviles en el área de la salud. Acad. J. (2015). ISBN 9781939982179
5. López, J.: Cloud Computing Analized from the Spanish Legal System (2013)
6. <http://Imms.gob.mx>
7. Sampieri, H., Fernández, C., Baptista, L.: Metodología de la Investigación, Editorial, Mc Graw Hill (2014)

**Part III**

**Advances in Data and Knowledge  
Management for Healthcare**

# Automatic Quantification of the Extracellular Matrix Degradation Produced by Tumor Cells

Nadia Brancati, Giuseppe De Pietro, Maria Frucci, Chiara Amoruso,  
Daniela Corda and Alessia Varone

**Abstract** Understanding the mechanisms of invasion of cancer cells into surrounding tissues is of primary importance for limiting tumor progression. The degradation of the extracellular matrix (ECM) and the consequent invasion of the surrounding tissue by tumor cells represent the first stage in the development and dissemination of metastasis. The quantification of such a degradation is thus an important parameter to evaluate the metastatic potential of tumor cells. Assessment of degradation is usually performed in *in vitro assays*, in which tumor cells are cultured on a gelatin (or other matrix)-coated dishes and the degraded gelatin areas under the tumor cells are visualized and quantified by fluorescent labelling. In this paper, we present an automatic method to quantify the ECM degradation through the feature analysis of the digital images, obtained from the *in vitro assays* and showing the tumor cells and the degraded gelatin areas. Differently from the existing methods of image analysis supporting biologists, our method does not require any interaction with the user providing quickly corrected and unbiased measures. Comparative results with a method frequently used by biologists, has been performed.

---

N. Brancati (✉) · G. De Pietro · M. Frucci

Institute for High Performance Computing and Networking,  
National Research Council of Italy (ICAR-CNR), Naples, Italy  
e-mail: nadia.brancati@cnr.it

G. De Pietro

e-mail: giuseppe.depietro@cnr.it

M. Frucci

e-mail: maria.frucci@cnr.it

C. Amoruso · D. Corda · A. Varone

Institute of Protein Biochemistry, National Research Council of Italy (IBP-CNR),  
Naples, Italy  
e-mail: c.amoruso@ibp.cnr.it

D. Corda

e-mail: d.corda@ibp.cnr.it

A. Varone

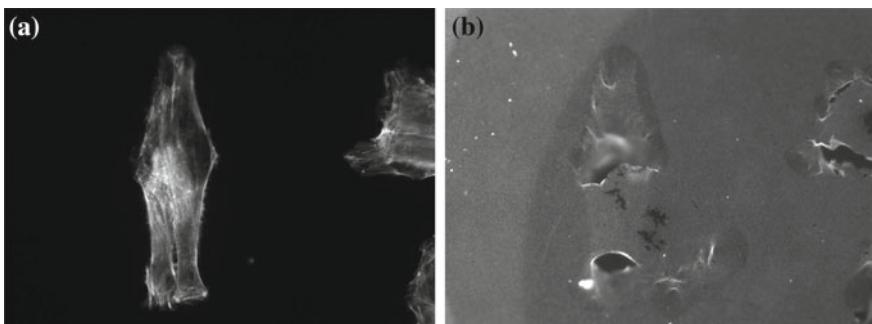
e-mail: a.varone@ibp.cnr.it

**Keywords** Extracellular matrix degradation • Binarization • Feature extraction

## 1 Introduction

The extracellular matrix (ECM) is a highly dynamic structure that is present in all tissues that continuously undergoes controlled remodeling; this is an essential process for development, wound healing and normal organ homeostasis. Many pathological conditions, such as inflammation, cardiovascular disease and tumor cells metastasis, arise when extracellular matrix remodeling becomes excessive or uncontrolled [12]. The analysis and quantification of the ECM degradation by cancer cells is thus a crucial feature to consider in order to determine their metastatic potential [4, 5, 20, 21]. To analyze the invasiveness of tumor cells by means of quantification of the ECM degradation, *in vitro* assays based on cells cultured on a gelatin-coated support, are often used [10]. Measures of the possible degradation of the gelatin matrix allow a quantification of the invasiveness of tumor cells. Different fluorescence substances are generally used to label with different colors the cells and the gelatin support. This process is required to allow the acquisition of two digital images, after a given period of observation of the assay of interest: one image includes only the tumor cells (in the following, cell image), while the second image represents the gelatin support (in the following, gelatin image) located under the tumor cells of the first image (see Fig. 1). The digital images are processed to perform the measure of the degradation areas for each tumor cell. Specifically, segmentation processes of the two acquired images are used to extract the regions of interest (cells and degradation areas, respectively); then, image analysis methods are performed to find both the correspondence between a cell and the relative degradation areas, and the computation of the area of these regions.

Many segmentation methods have been reported for the analysis of biological preparations [3, 8]. They differ in the type of measurements to be carried out [17, 19], in the type of cell images [2, 16] and in the complexity of these images



**Fig. 1** Example of an assay. **a** Cell image; **b** gelatin image

[18, 22]. In particular, for the types of images of interest in the present study, segmentation processes based on threshold algorithms are commonly adopted: they produce a binary image in which the foreground represents the regions of interest. The binarization processes and the quantification of the degradation areas are performed by means of standardized procedures based on available image processing software [1, 6, 13, 14].

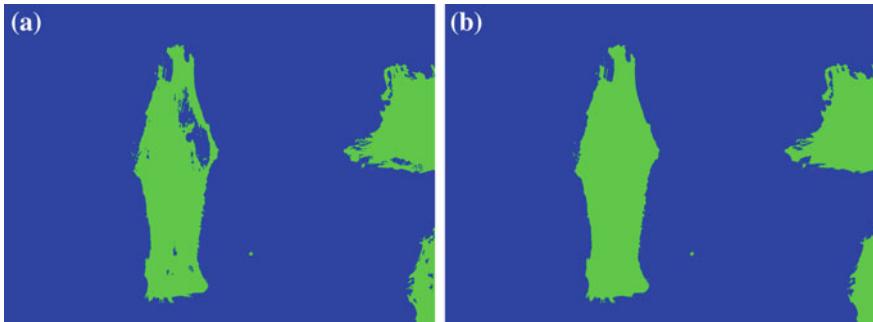
One of the most popular procedures used for this scope, is described in [13] and involves the following steps:

- a manual threshold setting for each binarization process;
- an indication of the feature to compute (for this context, the area);
- the choice of processes to eliminate the noise;
- the selection of the cells for which the degradation areas has to be calculated.

The main limitation of such a procedure is the necessary interaction with the user for the choice of the threshold values to apply in the binarization processes and the methods to remove noise from foreground. Such an interaction can produce arbitrary and subjective measures [7, 9]. In this paper, we propose a method to perform a fully automatic computation of the degradation areas, obtaining quickly results on the whole set of tumor cells under examination.

## 2 Method

The binarization of the cell image is performed using the faster version [11] of the Multi Otsu Threshold Algorithm [15]. This algorithm requires the definition of the following parameters: the range  $[m, M]$  of gray levels in which the threshold values should be computed and the number  $n$  of classes in which the image should be partitioned. Generally  $m = 0$  and  $M = 255$ ; in this work, the first parameter is automatically defined considering the minimum and maximum values of the gray levels associated to current image. On the basis of the value of the second parameter  $n$ , the algorithm provides  $n$  threshold values  $t_1 = m, t_2 \dots, t_n < M$ . The choice of the second parameter  $n$  depends on the type of the image to be segmented. Clearly, a binarization process implies a partition of the image in two classes (foreground and background) and for this,  $n$  should be set to 2 to generate a threshold value  $m < t_2 < M$ ; however, the partition performed on the basis of the value  $t_2$  might produce an under or over segmentation. For this, a value  $n > 2$  can be appropriate to generate more threshold values and use one of the threshold values or the combination of these values to obtain the correct partition of the image. To obtain the appropriate values of  $n$  for both binarization processes, a set of tests has been performed on a dataset of 30 assays, including the digital images of the tumor cells and gelatin supports. Accordingly, the appropriate number  $n$  of classes, for the cell images, has been chosen equal to 3. Thus, the algorithm provides three threshold values  $t_1 = m, t_2$  and  $t_3 < M$ ; the foreground of the cell image is represented from pixels with gray levels belonging to the range  $[t_2, M]$ , while the set of remaining pixels represents the background (see



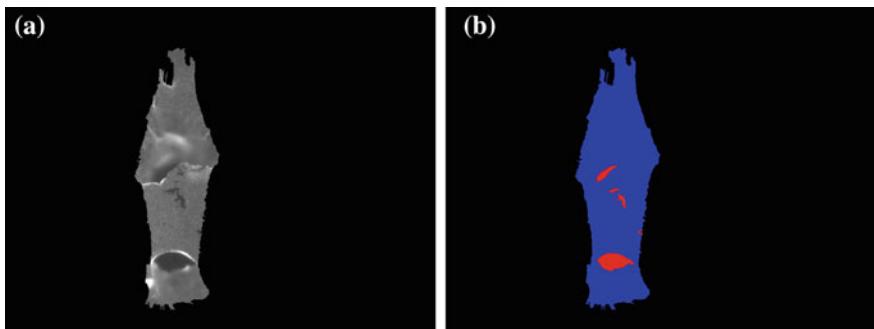
**Fig. 2** binarization of the image of Fig. 1. **a** Foreground in *green* and background in *blue*; **b** reduction of noise

Fig. 2a). To reduce the noise in the binarized image, the holes present in the regions of the foreground are filled (see Fig. 2b).

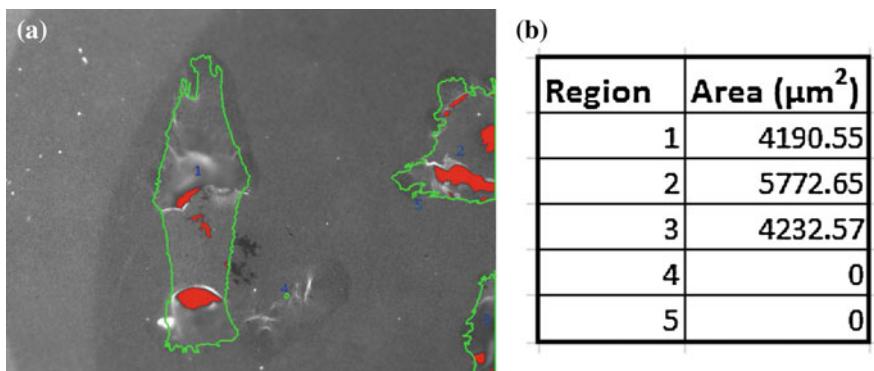
The foreground associated to the cell image is now used as a mask to binarize only the parts of gelatin image in which degradation areas could be detected. In particular, a connected component labeling is performed on the foreground of cell image so that any region of the foreground, representing a cell, is labeled by a single value. For each cell, a new image is built including only the part of the gelatin image corresponding to the selected cell (see Fig. 3a). The same threshold algorithm used for the cell image is applied on this new image, but the values  $m$  e  $M$  are computed only in the area masked by the current cell and  $n$  is experimentally set to 5 to generate five threshold values  $t_1 = m, \dots, t_5 < M$ . Since the degradation regions are the dark regions of the images (on the contrary, the cells are lighter than other regions in the relative image), pixels with low gray levels should be taken into account for the detection of the foreground. Precisely, the foreground of the current image is represented from pixels with gray levels belonging to the range  $[t_1, (t_2 + t_3)/2]$ ; the set of remaining pixels represents the background. To eliminate the noise on the found foreground, the same process applied to cell image is carried out (see Fig. 3b). Finally, the area of the foreground is computed.

As soon as the degradation areas of all cells are evaluated, the final result is shown to the user. More precisely, the contours of the cells are outlined in green, and the gelatin degradation areas are visualized in red (see Fig. 4a); moreover, the values of the total degradation area for each cell are given in a table (see Fig. 4b).

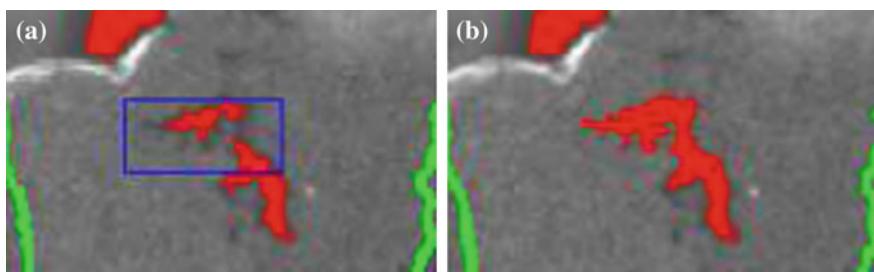
Since in some cases, several small degradation areas may not be detected, an optional step to improve the binarization is introduced. In particular, the user can draw a rectangle around the region of interest (ROI). The binarization, with  $n = 3$ , is performed on the ROI and the foreground is given by the set of pixels with value less or equal to  $t_2$ . An example of this improvement process is displayed in Fig. 5.



**Fig. 3** **a** Gelatin support related to the first cell; **b** binarization with reduction of noise: foreground in red and background in blue

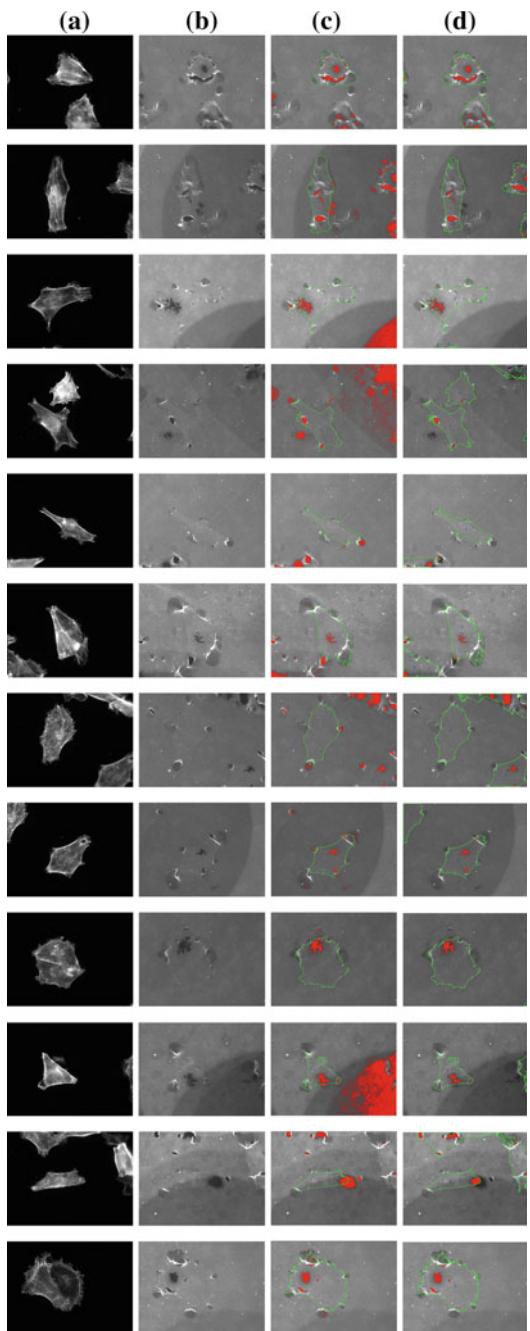


**Fig. 4** Result of the detection of the degradation areas. **a** Contours of the cells in green, and degradation regions in red; **b** measure of the degradation area



**Fig. 5** Example of semi-automatic improvement of the binarization on a magnified part of the Fig. 4a. **a** The blue rectangle is defined by the user to obtain a better detection of the degraded area in this part; **b** improved detection of degraded area

**Fig. 6** **a** Cells images; **b** gelatin images; **c** results of the procedure [13] for the cell outlined in green. Only the red degraded areas included into the selected cell are successively taken into account; **d** results of the proposed, automatic method



### 3 Results

To evaluate the advantages of the method presented in this study a comparison with the procedure proposed in [13] has been performed. For this, biologists were employed to measure the degradation areas for the dataset of 30 assays, by adopting the standard procedure proposed in [13] and applying, at the same time, our method without taking into account the optional step for the improvement of the binarization. Biologists evaluated the results and confirmed that our method produces generally a comparable or a more accurate detection of degraded areas, for a set of samples large enough. Some examples of the results obtained by applying the two methods are shown in Fig. 6, while comparisons of the quantitative results on the computation of the degradation areas for a given cell, are reported in Table 1. Since different results for the procedure [13] has been obtained by different users, the manual threshold values reported in the Table 1, represent the average of the values provided by biologists employed in the test.

Generally, the threshold values obtained manually for the binarization of the gelatin image are slightly higher than the values automatically computed. This implies that the degradation areas calculated with the procedure in [13] are always a little greater than the ones computed with the method we are reporting. A reason for this phenomenon is partly due to the incorrect human perception of contours of not well defined degraded areas and to the presence of small noisy regions produced by the over segmentation.

According to the overall evaluation of biologists on the use of our method, the advantages can be summarized as follows: (1) the measures can be obtained very

**Table 1** Results of the calculation of the degradation areas

Image in row	Manual threshold values		Automatic threshold values		Degradation area in $\mu\text{m}^2$	
	Cell	Gelatine	Cell	Gelatine	Procedure in [13]	Proposed method
1	15	80	46	77	7677.8	7155.2
2	15	80	42	72	5563.35	4190.55
3	25	90	37	80	6041.75	4442.75
4	30	90	40	69	2654.6	1968.2
5	25	100	45	103	173.55	109.2
6	25	90	41	80	2735.2	1851.2
7	35	80	43	70	289.9	117
8	30	70	44	67	3285.75	2883.4
9	45	60	40	57	7089.55	6268.6
10	45	60	46	65	4767.1	3784.3
11	60	70	46	69	3999.45	3804.45
12	40	90	35	81	6924.45	6002.1

rapidly and consistently; (2) no requirement for selecting a threshold values as for the procedure [13]; (3) reproducibility of the results, independently from the operators; (4) a more accurate identification of the degraded areas.

## 4 Conclusion

In this paper a fully automatic method for the computation of the ECM degradation areas produced by tumor cells is presented. The method is based on automatic binarization processes of the images obtained on assays of tumor cells cultured on gelatin-coated dishes. Comparisons with the standard semi-automatic procedures reported earlier [13], indicate that the detection and the quantification of the degradation areas are faster and their identification more accurate according to our method; moreover, the choice of parameters is completely unbiased as it is independent from the operators.

## References

1. Acerbi, I., Cassereau, L., Dean, I., Shi, Q., Au, A., Park, C., Chen, Y.Y., Liphardt, J., Hwang, E., Weaver, V.M.: Human breast cancer invasion and aggression correlates with ecm stiffening and immune cell infiltration. *Integr. Biol.* (2015)
2. Arslan, S., Ozyurek, E., Gunduz-Demir, C.: A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images. *Cytom. Part A* **85**(6), 480–490 (2014)
3. Brancati, N., Frucci, M., di Baja, G.S.: Image segmentation via iterative histogram thresholding and morphological features analysis. In: *Image Analysis and Recognition*, pp. 132–141. Springer (2008)
4. Bravo-Cordero, J.J., Hodgson, L., Condeelis, J.: Directed cell invasion and migration during metastasis. *Curr. Opinion Cell Biol.* **24**(2), 277–283 (2012)
5. Buccione, R., Caldieri, G., Ayala, I.: Invadopodia: specialized tumor cell structures for the focal degradation of the extracellular matrix. *Cancer Metastasis Rev.* **28**(1–2), 137–149 (2009)
6. Díaz, B.: Invadopodia detection and gelatin degradation assay. *Cancer* **3**(24) (2013)
7. Frazer, G.W., Fournier, R.A., Trofymow, J., Hall, R.J.: A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. *Agric. Forest Meteorol.* **109**(4), 249–263 (2001)
8. Frucci, M., Arcelli, C., Di Sanniti, G.: On the hierarchical assignment to the foreground of gray-level image subsets. *Int. J. Pattern Recognit. Artif. Intell.* **20**(06), 897–912 (2006)
9. Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F.: Review of methods for in situ leaf area index determination: part i. Theories, sensors and hemispherical photography. *Agric. Forest Meteorol.* **121**(1), 19–35 (2004)
10. Kramer, N., Walzl, A., Unger, C., Rosner, M., Krupitz, G., Hengstschläger, M., Dolznig, H.: In vitro cell migration and invasion assays. *Mutat. Res./Rev. Mutat. Res.* **752**(1), 10–24 (2013)
11. Liao, P.S., Chen, T.S., Chung, P.C.: A fast algorithm for multilevel thresholding. *J. Inf. Sci. Eng.* **17**(5), 713–727 (2001)
12. Lu, P., Takai, K., Weaver, V.M., Werb, Z.: Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor Perspect. Biol.* **3**(12), a005058 (2011)

13. Martin, K.H., Hayes, K.E., Walk, E.L., Ammer, A.G., Markwell, S.M., Weed, S.A.: Quantitative measurement of invadopodia-mediated extracellular matrix proteolysis in single and multicellular contexts. *J. Vis. Exp. JoVE* **66** (2012)
14. Martín-Villar, E., Borda-d'Agua, B., Carrasco-Ramírez, P., Renart, J., Parsons, M., Quintanilla, M., Jones, G.: Podoplanin mediates ecm degradation by squamous carcinoma cells through control of invadopodia stability. *Oncogene* (2014)
15. Otsu, N.: A threshold selection method from gray-level histograms. *Automatica* **11**(285–296), 23–27 (1975)
16. Saha, B., Saini, A., Ray, N., Greiner, R., Hugh, J., Tambasco, M.: A robust convergence index filter for breast cancer cell segmentation. In: 2014 IEEE International Conference on Image Processing (ICIP), pp. 922–926. IEEE (2014)
17. Seroussi, I., Veikherman, D., Ofer, N., Yehudai-Resheff, S., Keren, K.: Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes. *J. Microsc.* **247**(2), 137–146 (2012)
18. Su, H., Yin, Z., Huh, S., Kanade, T., Zhu, J.: Interactive cell segmentation based on active and semi-supervised learning (2015)
19. Wang, C.W., Fennell, D., Paul, I., Savage, K., Hamilton, P.: Robust automated tumour segmentation on histological and immunohistochemical tissue images. *PloS One* **6**(2), e15818 (2011)
20. Weaver, A.M.: Invadopodia: specialized cell structures for cancer invasion. *Clin. Exp. Metastasis* **23**(2), 97–105 (2006)
21. Yamaguchi, H., Wyckoff, J., Condeelis, J.: Cell migration in tumors. *Curr. Opin. Cell Biol.* **17**(5), 559–564 (2005)
22. Yin, Z., Su, H., Ker, E., Li, M., Li, H.: Cell-sensitive microscopy imaging for cell image segmentation. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014, pp. 41–48. Springer (2014)

# **Quantitative EEG and Virtual Reality to Support Post-stroke Rehabilitation at Home**

**Alfonso Mastropietro, Sara Arlati, Simona Mrakic-Sposta,  
Luca Fontana, Cristina Franchin, Matteo Malosio, Simone Pittaccio,  
Cristina Gramigna, Franco Molteni, Marco Sacco  
and Giovanna Rizzo**

**Abstract** Post-stroke rehabilitation has an enormous impact on health services worldwide because of the high prevalence of stroke, in continuous growth due to the progressive population aging. Systems for neuro-motor rehabilitation at home can help reduce the economic burden of long lasting treatment in chronic post-stroke patients; however the efficacy of these systems in providing a correct and effective rehabilitation should be established. From this point of view, coupling home rehabilitation systems with quantitative EEG methodologies for objectively characterizing patients' cerebral activity could be useful for the clinician to optimize the rehabilitation protocol and assess its efficacy. Moreover, the use of virtual/augmented reality technologies could assist the patients during unsupervised rehabilitation by providing an empathic feedback to improve their adherence to the treatment. These two aspects were studied and implemented in RIPRENDO@home, a multidisciplinary project, aimed to develop an integrated technological platform oriented to home neurorehabilitation for stroke patients.

**Keywords** Neuro-rehabilitation • Stroke • EEG • Virtual reality

---

A. Mastropietro (✉) · S. Mrakic-Sposta · C. Franchin · G. Rizzo  
Istituto di Bioimmagini e Fisiologia Molecolare (IBFM), Consiglio Nazionale delle Ricerche,  
via Fratelli Cervi 93, 20090 Segrate, MI, Italy  
e-mail: alfonso.mastropietro@ibfm.cnr.it

S. Arlati · L. Fontana · M. Malosio · M. Sacco  
Istituto di Tecnologie Industriali e Automazione (ITIA), Consiglio Nazionale delle Ricerche,  
via Bassini 15, 20133 Milan, Italy

S. Pittaccio  
Istituto per l'Energetica e le Interfasi (IENI), Consiglio Nazionale delle Ricerche,  
Corso Promessi Sposi 29, 23900 Lecco, Italy

C. Gramigna · F. Molteni  
Villa Beretta, Ospedale Valduce, via Nazario Sauro 17, 23845 Costa Masnaga, LC, Italy

## 1 Introduction

Stroke is one of the most important cause of long-term disability and its effects may involve the physical, psychological, social, and financial spheres for both patients and their families. In the developed countries, stroke is the second leading cause of mortality [1, 2] and the continuous ageing of the population is an important cause of soaring stroke incidence.

Stroke represents an expensive economic burden for the health systems; its costs are up to 3 % of all yearly health expenditure in industrialized countries [3]. Neurological damages due to cerebrovascular diseases are a worldwide relevant issue and rehabilitative interventions should be implemented in order to improve patients' conditions and reduce the socio-economic impact of the stroke [3].

Approximately two thirds of the patients affected by cerebrovascular diseases require rehabilitation and most of them present residual and disabling long-term deficits due to impaired motor function. In order to increase the patients' life quality and to reduce the economic burden, transferring the proposed therapy to the living environment of the patient can be a smart strategy. For this reason home-based neuro-rehabilitation systems have been recently proposed [4].

In this scenario, the RIPRENDO@home project was conceived by a multidisciplinary team of researchers working together in the field of neurorehabilitation at the new center of the Italian National Council of Research, located in Lecco, Lombardy. The project proposes an integrated technological platform oriented to the home neurorehabilitation of the upper limb in stroke patients. The platform is composed by (i) active devices for arm motion, dynamic repositioning orthoses with pseudoelastic properties and multiparameter monitoring system, (ii) techniques and methods for a quantitative functional patients' characterization, along with (iii) virtual/augmented reality technologies for patient engagement and adherence to the treatment. Among the different aspects tackled in the project, two of them are specifically devoted to improve and evaluate the effectiveness of home rehabilitation. Specifically, the automatic extraction of quantitative indices from electroencephalographic (EEG) signals, related to the cerebral activity of the patient, could objectively support monitoring of the rehabilitation protocol and assessing of the efficacy of the rehabilitation itself; the virtual/augmented reality technologies assist the patients during the unsupervised rehabilitation at home, by providing an empathic feedback aimed at improving his/her adherence to the treatment.

This work is aimed at presenting technologies and methods implemented in the RIPRENDO@home platform for introducing EEG analysis ad virtual/augmented reality in the framework of home-based rehabilitation. The paper is structured as follows: Sect. 2 describes the developed methodologies for the functional characterization of the patient's cerebral activity by quantitative electroencephalography (qEEG); Sect. 3 illustrates the developed "virtual reality" environment; Sect. 4 presents conclusions, future applications and challenges.

## 2 Quantitative EEG for Cerebral Activity Characterization

EEG reflects the activity of the cortical pyramidal cells and it is very sensitive in detecting the abnormalities of cerebral rhythms that are typical of stroke.

In particular, it is well established that quantitative EEG (qEEG) indices, based on the relationship between the power of slow and fast activity or measuring brain activity symmetry, as estimated by resting-state EEG power spectrum analysis, are reliable markers to characterize the brain status [5] and can be employed to improve patient management during ischemic stroke [6] and to predict clinical efficacy of rehabilitation even in the chronic stage [7].

Moreover, brain neuronal responses related to voluntary movements are known to produce EEG power modulations in alpha (8–12 Hz) and beta (12.5–30 Hz) frequency bands in the motor areas [8]. In the specific context of neuro-rehabilitation, recent studies explored the effects of rehabilitations using robotic devices, reporting alpha and beta desynchronization in the motor areas during sensorimotor processing [9, 10].

Starting from this neurophysiological knowledge, in RIPREND@home we developed both EEG resting state spectral analysis and EEG movement-related time-frequency analysis to characterize the electrical cortical activity of the patients and assess its evolution after the rehabilitation treatments.

### 2.1 QEEG in Resting-State Conditions

In RIPREND@home, the resting state EEG analysis was based on three minutes of EEG acquisition, recorded with the subject in supine position with eyes closed. A Synamps 2/RT EEG system (Neuroscan) was used to record cerebral activity. The system has a 70-channels amplifier system, consisting of 64 monopolar, 4 bipolar, and 2 high-level input channels. All electrodes were placed according to the international 10–20 electrode placement standard. EEG signals were acquired at a sampling frequency of 1 kHz.

Data pre-processing and filtering were performed in Matlab (The Mathworks, MA) using EEGLAB (<http://www.sccn.ucsd.edu/eeglab/index.html>) [11] and homemade scripts. Signals were resampled at 500 Hz and low-pass filtered at 45 Hz (resting-state EEG). Artifacts were removed using an Independent Component Analysis (ICA) [12]. Frequency bands considered for this study were delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz) and beta (13–30 Hz).

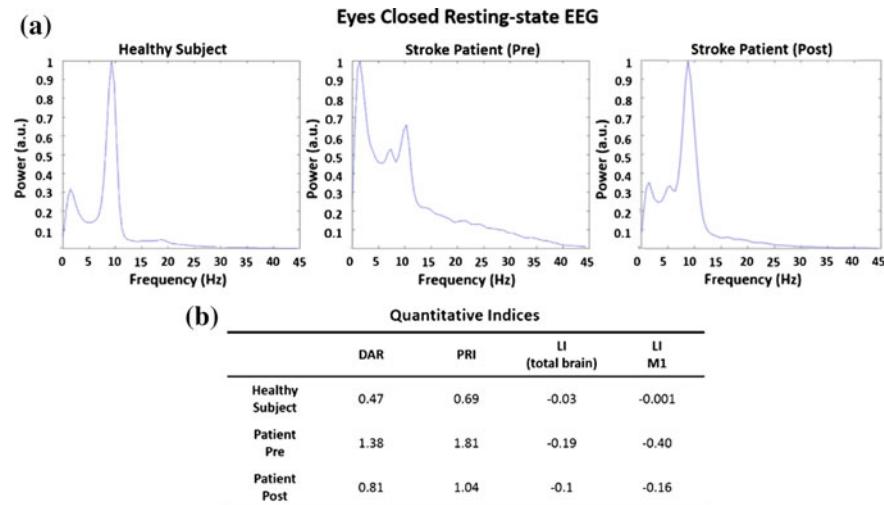
Power spectral density was estimated for each channel using the Welch's method. From the average power spectra, the mean power was computed across the frequency bands. The resulting mean bandpower values were used to automatically calculate the following quantitative indices:

1. **Delta/Alpha Ratio (DAR)**: defined as the ratio of delta to alpha absolute power.
2. **Power Ratio Index (PRI)**: the ratio of “slow” to “fast” activity defined as the ratio of delta-plus-theta to alpha-plus-beta absolute power.
3. **Laterality Index (LI)**:

$$LI = \frac{AP_{ipsi} - AP_{contra}}{AP_{ipsi} + AP_{contra}} \quad (1)$$

where  $AP_{ipsi}$  and  $AP_{contra}$  are the absolute powers of corresponding ipsilateral and the contralateral channels (electrodes), respectively.

As an example, in Fig. 1a, power spectral densities in a healthy subject and a stroke patient (pre and post upper-limb rehabilitation, 1 month) are shown, as obtained by our procedure. With eyes closed, the EEG spectrum of the awake healthy subject is characterized by a high power in alpha frequency band. Conversely, this particular patient exhibits a higher slow activity (delta and theta) before the rehabilitation treatment. After the rehabilitation, the power spectrum of patient displayed a more evident alpha power component, which is more similar to the healthy subject. In Fig. 1b the quantitative indices extracted from the power spectra are highlighted. In the healthy subject the brain activity is almost completely symmetric ( $LI \approx 0$ ). The patient’s brain activity is more asymmetric as indicated by the LI values in both total brain and M1 (primary motor area). It is reasonable to



**Fig. 1** An example of resting-state EEG analysis. **a** Power spectra of a healthy subject and a patient (pre and post rehabilitation) acquired with eyes closed. **b** Quantitative indices are displayed in the same subjects. DAR, PRI and LI (total brain) where calculated as a mean value over all the brain electrodes indices. LI (M1) was calculated on the electrodes corresponding to the primary motor cortex M1 (C4, C6 right and C3, C5 left)

believe that the rehabilitation treatment had a role in partially compensating this asymmetry (LI decreases). In a similar way, DAR and PRI indices were found to decrease after the treatment in this patient, and became closer to the healthy subject's values.

## 2.2 QEEG During Movement

EEG acquisition during movement were performed with patient in supine position and eyes open. The same system described in the previous paragraph was used. The patient was asked to carry out flexion extension movements of the elbow (30 movements, 10 s of rest after each movement). The duration of the protocol was about 5 min. Four surface electromyography (EMG) bipolar electrodes were used as movement trigger and located on different upper limb muscle (biceps, brachioradialis, anterior deltoid, triceps).

In contrast to the resting-state processing, EEG data were band-pass filtered (2.5–45 Hz) and the data were epoched to distinguish each movement. A time-frequency analysis was performed in order to obtain the event-related spectral perturbation (ERSP) which measures the mean event-related changes in the power spectrum in each data channel [11, 13]. In order to calculate ERSP, the power spectrum over a sliding latency window must be computed and then averaged across data trials. ERSP is defined as:

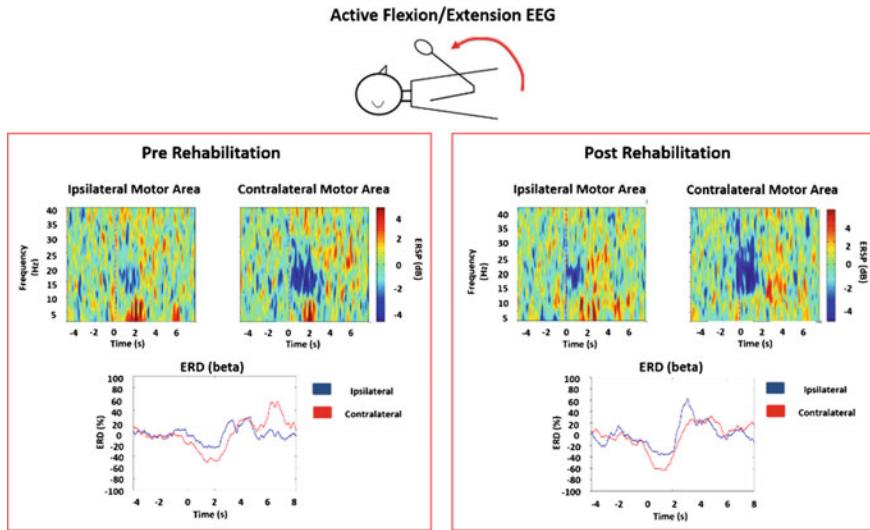
$$ERSP = \frac{1}{n} \sum_{k=1}^n |F_k(f, t)|^2 \quad (2)$$

where  $f$  and  $t$  are the frequency and the time respectively,  $n$  is the number of epochs,  $F_k(f, t)$  is the spectral estimate of trial  $k$  at frequency  $f$  and time  $t$  and was computed using a sinusoidal wavelet (short-time DFT) transform. ERSP can be viewed as a generalization of the event-related desynchronization (ERD) [14], where ERD is defined as:

$$ERD = \frac{BP_{mov} - BP_{rest}}{BP_{rest}} \quad (3)$$

where  $BP_{mov}$  is the power within the frequency band of interest (alpha or beta) during the activity period, while  $BP_{rest}$  is evaluated in the baseline EEG signal before the movement starts.

In Fig. 2, an example of time frequency analysis of EEG acquired on a stroke patient during the flexion extension movement of the elbow is displayed. The graphs at the bottom display the mean ERD in beta band in both ipsi and contralateral M1 areas. Before the rehabilitation a residual weak desynchronization in beta band is appreciable in contralateral motor area and in the ipsilateral area. After the rehabilitative treatment, in this subject, the activation of the contralateral M1 is



**Fig. 2** An Example of time-frequency analysis of the EEG signal acquired during active elbow flexion/extension tasks. ERSP of a stroke patient in contralateral and ipsilateral primary motor cortex is displayed (*upper panels*). Blue areas are indicative of brain desynchronization while red areas are related to resynchronization. ERD in beta band are displayed in the *bottom panels*

more evident as a desynchronization in beta band and the difference between ipsilateral and contralateral area is increased as well.

### 3 Virtual Reality for Upper Limb Rehabilitation

During the last few years, various systems that exploit virtual or augmented reality (VR/AR) were developed with the aim of entertaining and/or motivating patients during their rehabilitation path. The need to provide patients with engaging scenarios and supporting mechanisms during the rehabilitative exercises arises from the fact that therapy often requires high-frequency repetition of the same movements to be effective [15]. The repetitiveness of the same tasks, and the lack of immediate remarkable improvements, may lead the patients to abandon rehabilitation—especially when they are expected to be doing unsupervised at-home rehabilitation [16].

In this scenario, VR/AR applications appear as good means to increase patients' motivation and adherence. Game-based elements (scores, changing environments, rewards, etc.) act as key elements to boost users' motivation and have been largely employed in many VR systems for rehabilitation [17–19]. User satisfaction questionnaires often reported a positive opinion of the patients toward this kind of alternative therapy [20], although no study seems to have been conducted yet

including an appropriate follow-up period [18]. Thus, it is not possible to determine if boredom or other issues emerge for longer periods than the trial (which typically lasts from 4 to 12 weeks).

Another important aspect linked to the use of digital environments for rehabilitation is the possibility to provide patients with multisensory feedback, in order to improve their performance by increasing their body awareness [21–23]. This aspect not only takes into account specific patient's desires [24], but also allows creating home-based rehabilitation systems, able to identify wrong behaviors automatically and signal them to the patients using visual, auditory or haptic cues.

Within the Riprendo@HOME project, we followed these indications to develop a VR application, called REAPP [25–27]. REAPP was not designed as a game, but with the aim of providing different kinds of feedback, appropriately weighted on the patient's cognitive capabilities and mood, to improve patient's body and movement awareness and make him/her more conscious and involved in the rehabilitation path.

### **3.1 System Architecture**

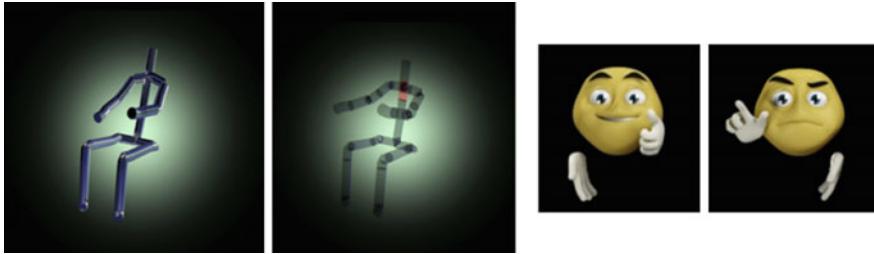
Since the REAPP system was conceived for domestic use, one of the main constraints was to keep the costs low. Thus, the minimum requirements for this VR application dedicated to stroke patients are a Windows-8 PC with a screen and a low-cost tracking sensor: the Microsoft Kinect v2.

In order to accommodate user-specific requirements, REAPP implements several widgets that provide different types of feedback, which can be shown or inactivated, depending on the patient's characteristics or mood [26].

The first REAPP key widget is the patient's avatar, which provides a visual feed-back about the positions of all body segments during the execution of a certain rehabilitation task. The 3D avatar is rendered in real-time on the patient-dedicated graphical user interface (GUI), using the data coming from the Kinect sensor. The choice of using an avatar (rather than, for instance, a video camera) offers the possibility to change the point-of-view either focusing patients' attention onto a specific body part, or letting them choose their favorite viewpoint.

Moreover, the avatar allows giving the patients an immediate visual feedback when a certain body part is out of the optimal range or trajectory, through the red blinking of the corresponding virtual body segment. The constraints which determine the optimal ranges, in which the single body parts should stay during specific tasks, must be defined by the therapist in a medical examination, during which the patient's capabilities and limitations are assessed.

The other key widget of REAPP is the virtual assistant (VA) (Fig. 3). It substitutes the human therapist during the unattended home sessions, taking care of all of the instructions and messages that have to be delivered to the patient. The one-to-one interaction between the patient and therapist is a crucial element for patient's recovery in traditional rehabilitation [28, 29], and the positive aspects of



**Fig. 3** The patient avatar (blinking in red to signal an error) and the virtual assistant

this relationship ought not to be lost in case of virtual-assisted domestic rehabilitation sessions [24]. To try and keep the advantages of the human-to-human interaction, avoiding any other kind of bias (e.g. race, gender, etc.), the VA was designed as an animated smiley-like character that emphasizes the messages both with the voice tone and the hand gesture (i.e. thumb-up and wink for the “congratulation” message) [30].

In addition, the VA asks the patients about their mood and their physical condition with the twofold aim of acquiring data for the human therapist (that can re-examine all the session data) and re-creating what normally happens at the beginning and at the end of a traditional therapy session. During the exercises, instead, the VA is not always present on the patient GUI but appears when there is the need to correct the patient after a repeatedly wrong execution of the motor task (for which the visual clue from the avatar was not enough) or to encourage him/her because he/she is performing well. The temporization between two subsequent appearances of the VA can be varied in order not to overwhelm the patient with too many pieces of information in a little time.

Besides these two widgets, the patient GUI can be enriched with other widgets that are exercise-specific and were devised to help the patient accomplish the exercise in the best possible way (e.g. target/direction indicators). Additional widgets and exercises, even associated with different robotic devices [26, 27], can be easily added to REAPP, since the entire software has been designed to be modular and flexible, so that the same core can be adopted by different patients, providing each one with his/her own customized platform.

## 4 Conclusions

In this paper, two methodologies to support the home-based rehabilitation treatment in chronic ischemic patients, which are parts of the RIPRENDÒ@home project, were described; (a) quantitative EEG for brain characterization activity and evaluation of the treatment and (b) VR environment to support the rehabilitation of the upper limb.

There is a need for quantitative and objective measures in the clinical practice to assess the efficacy of rehabilitation. For this reason, EEG and VR methods can help the caregivers to adapt the treatment to the patient's response and to assess day by day the patient's progress and his/her status.

The methodologies presented in this paper, even if useful and robust, were tested on a small number of subjects and for a limited period of time, so no information about the rehabilitation effects can be drawn yet. However some interesting topics and applications have emerged for future works and will be soon tested in a clinical environment, thanks to the strong collaboration of the clinical centers that work on neurorehabilitation in the Lecco's area.

More in detail, further researches will be conducted to determine if REAPP, as it is, can be a good mean to promote rehabilitation at-home or if, instead, minor (e.g. better customization of the avatar) or major changes (e.g. use of immersive environment or AR) are required to provide a more appropriate users' experience. Within this context, EEG could be used not only to provide a prediction on the outcome of the treatment but also to investigate how much the patient is involved during the virtual rehabilitation training. EEG and VR can be integrated in more complex tele-health systems or decision support systems to improve the patient management and the rehabilitative process directly at home reducing the economic burden of the health systems. Patients at home should be more motivated to perform the proposed treatment and this can improve the efficacy of the therapy and the life quality of the subjects and their families.

**Acknowledgements** The authors wish to thank Dr. Fabio Rastelli for his useful contribution in EEG acquisitions and Dr. Stefano Mottura, Dr. Claudia Redaelli and Dr. Andrea Zangiacomi for their contribution to the REAPP development. The authors want to thank the Scientific Institute IRCCS Eugenio Medea (Bosisio Parini, Italy) and the Lombardy Cluster "Technologies for Living Environments" for supporting the project activity.

The work was performed within the RIPRENDÒ@Home Project, regional research project funded inside the Framework Agreement between Regione Lombardia and National Research Council, D.G.R. n. 3728-11/07/2012.

## References

1. World Health Organization: The Global Burden of Disease: 2004 update. Update, vol. 2010 (2008). doi:[10.1038/npp.2011.85](https://doi.org/10.1038/npp.2011.85)
2. Millán, M., Dávalos, A.: The need for new therapies for acute ischaemic stroke. *Cerebrovasc. Dis.* **22**, 3–9 (2006)
3. WHO: Neurological disorders: a public health approach. *Neurol. Disord. Public Health Challeng.*, 41–176 (2006)
4. Johnson, M.J., Feng, X., Johnson, L.M., Winters, J.M.: Potential of a suite of robot/computer-assisted motivating systems for personalized, home-based, stroke rehabilitation. *J. Neuroeng. Rehabil.* **4**, 6 (2007)
5. Corsi-Cabrera, M., Galindo-Vilchis, L., del-Río-Portilla, Y., Arce, C., Ramos-Loyo, J.: Within-subject reliability and inter-session stability of EEG power and coherent activity in women evaluated monthly over nine months. *Clin. Neurophysiol.* **118**, 9–21 (2007)

6. Finnigan, S., van Putten, M.J.A.M.: EEG in ischaemic stroke: quantitative EEG can uniquely inform (sub-)acute prognoses and clinical management. *Clin. Neurophysiol.* **124**, 10–19 (2013)
7. Leon-Carrion, J., Martin-Rodriguez, J.F., Damas-Lopez, J., Barroso y Martin, J.M., Dominguez-Morales, M.R.: Delta-alpha ratio correlates with level of recovery after neurorehabilitation in patients with acquired brain injury. *Clin. Neurophysiol.* **120**, 1039–1045 (2009)
8. Pfurtscheller, G.: Lopes da Silva, F.H.: Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* **110**, 1842–1857 (1999)
9. Sale, P., Ingarinato, F., Del Percio, C., Lizio, R., Babiloni, C., Foti, C., Franceschini, M.: Electroencephalographic markers of robot-aided therapy in stroke patients for the evaluation of upper limb rehabilitation. *Int. J. Rehabil. Res. Int. Zeitschrift für Rehabil. Rev. Int. Rech. Réadaptation.* **38**, 294–305 (2015)
10. Gandolfi, M., Formaggio, E., Geroni, C., Storti, S.F., Boscolo Galazzo, I., Waldner, A., Manganotti, P., Smania, N.: Electroencephalographic changes of brain oscillatory activity after upper limb somatic sensation training in a patient with somatosensory deficit after stroke. *Clin. EEG Neurosci.* **46**, 347–352 (2015)
11. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004)
12. Iriarte, J., Urrestarazu, E., Valencia, M., Alegre, M., Malanda, A., Viteri, C., Artieda, J.: Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study. *J. Clin. Neurophysiol.* **20**, 249–257 (2003)
13. Makeig, S.: Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr. Clin. Neurophysiol.* **86**, 283–293 (1993)
14. Pfurtscheller, G.: Functional brain imaging based on ERD/ERS. *Vis. Res.* **41**, 1257–1260 (2001)
15. Wolf, S.L., Blanton, S., Baer, H., Breshears, J., Butler, A.J.: Repetitive task practice: a critical review of constraint-induced movement therapy in stroke. *Neurologist* **8**, 325–338 (2002)
16. Johnson, M.J., Johnson, L.M., Ramachandran, B., Winters, J.M., Kosasih, J.B.: Robotic Systems that Rehabilitate as well as Motivate: Three Strategies for Motivating Impaired Arm Use. In: The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. pp. 254–259. IEEE (2006)
17. Shah, N., Basteris, A., Amirabdollahian, F.: Design parameters in multimodal games for rehabilitation. *Games Health J.* **3**, 13–20 (2014)
18. Flores, E., Tobon, G., Cavallaro, E., Cavallaro, F.I., Perry, J.C., Keller, T.: Improving patient motivation in game development for motor deficit rehabilitation. In: Proceedings 2008 International Conference on Advances in Computer Entertainment Technology—ACE '08, vol. 7, p. 381 (2008)
19. Bayón-Calatayud, M., Peri, E., Fernández Nistal, F., Duff, M., Nieto-Escámez, F., Lange, B., Koenig, S.: Virtual rehabilitation. In: Emerging Therapies in Neurorehabilitation II, pp. 1–11 (2015)
20. Pastor, I., Hayes, H.A., Bamberg, S.J.M.: A feasibility study of an upper limb rehabilitation system using Kinect and computer games. In: Conference on Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1286–1289 (2012)
21. M. F. Levin, J. E. Deutsch, M. Kafri, and D. G. Liebermann, “Validity of Virtual Reality Environments for Sensorimotor Rehabilitation,” in *Virtual Reality for Physical and Motor Rehabilitation*, Springer, 2014, pp. 95–118
22. Parker, J., Mawson, S., Mountain, G., Nasr, N., Zheng, H.: Stroke patients’ utilisation of extrinsic feedback from computer-based technology in the home: a multiple case study realistic evaluation. *BMC Med. Inf. Decis. Mak.* **14**, 46 (2014)
23. Kiper, P., Agostini, M., Luque-Moreno, C., Tonin, P., Turolla, A.: Reinforced feedback in virtual environment for rehabilitation of upper extremity dysfunction after stroke: preliminary data from a randomized controlled trial. *Biomed Res. Int.* **2014**, 752128 (2014)

24. Zangiacomi, A., Redaelli, C., Valentini, F., Bernardelli, G.: Design of interaction in a virtual environment for post-stroke rehabilitation: a cognitive perspective. In: 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom). pp. 167–172. IEEE (2014)
25. Mottura, S., Arlati, S., Fontana, L., Sacco, M.: Enhancing awareness and personification by virtual reality and multimedia means in post-stroke patients during rehabilitation. In: 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom). pp. 179–184. IEEE (2014)
26. Mottura, S., Fontana, L., Arlati, S., Zangiacomi, A., Redaelli, C., Sacco, M.: A virtual reality system for strengthening awareness and participation in rehabilitation for post-stroke patients. *J. Multimodal User Interfaces.* **9**, 341–351 (2015)
27. S. Mottura, L. Fontana, S. Arlati, C. Redaelli, A. Zangiacomi, and M. Sacco, Focus on Patient in Virtual Reality-Assisted Rehabilitation in Virtual Reality Enhanced Robotic Systems for Disability Rehabilitation, p. 85 (2016)
28. Hall, A.M., Ferreira, P.H., Maher, C.G., Latimer, J., Ferreira, M.L.: The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Phys. Ther.* **90**, 1099–1110 (2010)
29. Maclean, N., Pound, P.: A critical review of the concept of patient motivation in the literature on physical rehabilitation. *Soc. Sci. Med.* **50**, 495–506 (2000)
30. Bates, J.: The role of emotion in believable agents. *Commun. ACM* **37**, 122–125 (1994)

# Towards a Sustainable Solution for Collaborative Healthcare Research

Nikos Karacapilidis and George Potamias

**Abstract** This paper describes a novel web-based environment that empowers healthcare research communities to efficiently and effectively collaborate thanks to reliable and user-friendly access to integrated and interoperable resources of different types. The proposed solution enables heterogeneous data and knowledge sources, as well as healthcare-related processing methodologies and tools, to be wrapped through appropriate web services. The paper sketches the motivation behind the proposed solution, discusses its architecture and types of services to be integrated, and comments on the foreseen advancements in healthcare research from a technological, a collaborative and an organizational perspective.

**Keywords** Healthcare • Knowledge management • Decision making • Collaboration • Virtual research environment

## 1 Introduction

Collaborative research environments require user-friendly solutions that mask the overall complexity of the problem, allowing stakeholders to easily handle big volumes of data and providing them with meaningful recommendations upon which they can base their decisions [1]. This is even more imperative in healthcare research, which is characterized by high levels of inter-disciplinarity, data-intensiveness and cognitive complexity. In such settings, knowledge management, sense making and decision making issues are heavily associated with voluminous multi-faceted data residing in heterogeneous resources, and need appropriate treatment [2, 3].

---

N. Karacapilidis (✉)  
IMIS Lab, MEAD, University of Patras, 26504 Rio Patras, Greece  
e-mail: nikos@mech.upatras.gr

G. Potamias  
Institute of Computer Science, FORTH, 71110 Heraklion, Greece

In line with the above, this paper proposes a sustainable virtual research environment that empowers healthcare research communities to efficiently and effectively collaborate thanks to reliable and user-friendly access to integrated and interoperable resources of different types. The proposed environment is an innovative web platform, where heterogeneous data and knowledge sources—as well as respective processing methodologies and tools—are wrapped through appropriate web services, while integration is performed on a service level. Exploiting and advancing the outcomes of the *DICODE* project [4], the proposed platform enables the seamless integration of heterogeneous services and ensures their interoperability from a technical, conceptual and user interface point of view.

As discussed in the next sections, our overall approach builds on the synergy between human and machine reasoning capabilities to tame information overload and cognitive complexity in healthcare research settings. It offers an innovative solution that improves the quality of the outcome of a collaboration process, while enabling users to be more productive and focus on creative activities. It provides researchers with the required capacity to easily assemble their own working environment using their preferred tools, as well as to exploit the wealth of available resources and competences.

## 2 Concept and Motivation

### 2.1 *The Current Situation in Conducting Healthcare Research*

The motivation that triggers and guides the development of the proposed collaborative research environment lies in the fact that contemporary healthcare research is heavily influenced by the current flood of digital content. The generated raw data are so overwhelming that healthcare researchers and stakeholders are often at a loss to even know where to begin to make sense of them. This is today a common situation in most scientific fields, where research employs a range of disciplines to form, acquire and synthesize the appropriate levels of evidence for targeted research hypotheses.

Furthermore, the ever-increasing rate in the delivery of research results narrows the utilitarian potential of accumulated knowledge. This is more intense in domains where the investigated and identified associations among intermediate results lack supporting evidence, making their translation into reliable recommendations and decisions unfulfilled. This is due to the fragmentation of healthcare research, which comprises: (a) *intra-discipline fragmentation*—the same research hypothesis is explored under different experimental set-ups and protocols (e.g., by using different platforms or data analysis methods); (b) *inter-discipline fragmentation*—the same research hypothesis is explored on different levels or from different perspectives. In both cases, research results and findings, for the same research question, remain

unlinked. In such a fragmented context, researchers become ‘isolated’; they are drowning in a flooded, even unlinked, digital pool of research results to support not only their findings but also the research question itself.

As a consequence, healthcare researchers often fail to make the right decisions about their research quests, as well as to get the most out of their research, mainly due to failure in knowledge coupling; that is, they miss the assimilation and reliable amalgamation of interdisciplinary knowledge constructs. This situation results in over-reliance of the unaided human mind to recall and organize all the relevant details. Consequently, collection, sharing, analysis and interpretation of large quantities of available data and knowledge (often referred as the Big Data challenge) do not only constitute foundational requirements, but also highly critical challenges in the field of healthcare research.

## 2.2 *Towards a New Research Ecosystem*

The proposed collaborative healthcare research (CoHeRe) platform aims to bridge a number of gaps that exist in interdisciplinary healthcare research communities, such as the *data gap* (diverse data sources to be linked), the *semantic gap* (multiple vocabularies and ontologies to be merged), the *knowledge gap* (heterogeneous, fast growing and fragmented knowledge resources to be associated), and the *collaboration gap* (differences in interests, objectives and methodologies of diverse stakeholders). The development of the proposed platform is based on open standards and custom web technology; this allows an easy extension of the platform by using and adapting existing resources (i.e. data resources and data analysis tools), or developing new ones to cover the needs of related contexts.

To meet the emerging and continuously increasing needs of healthcare research communities, we adopt a component-based approach where each component encapsulates a particular functionality (and related data) and is associated with a *web service*. The proposed solution distinguishes three layers of components—namely, the *data layer*, the *knowledge layer* and the *collaboration layer*—and offers a flexible, scalable and customizable information and computation infrastructure that: (i) provides healthcare research communities with the whole spectrum of commodities needed to accomplish their tasks and goals; (ii) exploits the competences of researchers and stakeholders by incorporating the underlying collective intelligence; (iii) enables controlled sharing of research outcomes and assures data provenance. Moreover, the proposed solution pays much attention to usability and ease-of-use issues, aiming to enable users without any particular programming expertise to use the proposed platform.

CoHeRe targets and promotes the maximum sharing and reuse of existing resources, including datasets, services, computing power and hosting machines, aiming to manage the scale and complexity of contemporary healthcare research processes through a sustainable solution. Appropriate resource management services enable users assemble the components they need in their research workflow

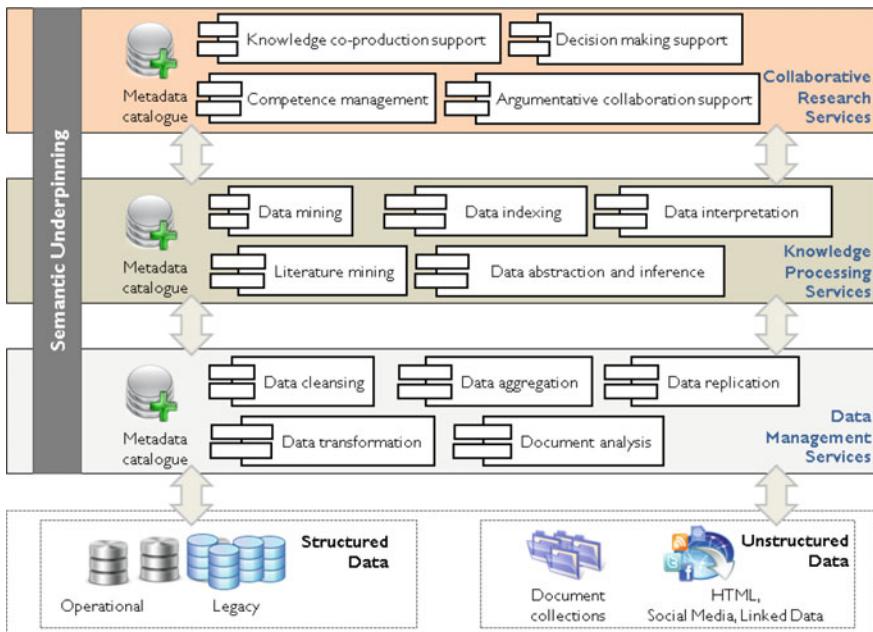
without worrying about implementation details and acquisition of necessary computational resources.

### 3 The Proposed Solution

CoHeRe provides a sustainable virtual research environment that empowers interdisciplinary healthcare research communities and stakeholders to efficiently and effectively collaborate thanks to reliable and user-friendly access to integrated and interoperable resources of different types. The proposed solution is an innovative web platform, where heterogeneous (both in their formats and content) data and knowledge sources, as well as respective processing methodologies and tools, are wrapped through appropriate web services and integration is performed on a service level. The CoHeRe solution does not pose any restrictions on the back-end technology. The platform enables the seamless integration of heterogeneous services and ensures their interoperability from a technical, conceptual and user interface point of view. Semantics techniques are highly exploited to define an ontological framework for capturing and representing the diverse stakeholders and services perspectives. Users are able to easily customize the platform through the proper assembly of web services and associated data resources that suit to their needs.

Through novel virtual workspaces, CoHeRe enables healthcare researchers and stakeholders to set up and engage in a highly interactive and collaborative process of discovery and evidence-based documentation of research indications and findings, where they can easily: decide about which data repositories should be considered; trigger and parameterize the associated data and knowledge processing mechanisms; explore and assess their discovery patterns; argue about the weaknesses and strengths of the identified patterns, assign evidential supports to them and rank putative hypotheses; control the complexity of the research output; set up new iterations of the data and knowledge processing procedures by defining other analytic alternatives or considering alternative services and data sources; exploit the wealth of available resources and expertise through intelligent resource and competence management services.

CoHeRe provides an innovative service-driven research environment that enables insightful collaboration among healthcare stakeholders thanks to a proper exploitation of the synergy between human and machine reasoning capabilities. This environment ensures easy access to integrated resources and streamlines diverse research activities. It is composed by an interoperable set of web services (Fig. 1), which meet the underlying requirements of capturing, delivering and analyzing pertinent information, and can be orchestrated through scientific workflows. The proposed solution is able to turn information growth into knowledge growth; it improves the quality of the outcome of collaborative healthcare research activities, while enabling users to be more productive and focus on creative activities.



**Fig. 1** The CoHeRe platform architecture

The CoHeRe platform seamlessly integrates the following types of services:

- *Data management services*, which enable the targeted discovery, capturing, archiving, sharing and processing of tractable large scale data existing in diverse data sources and formats. Much attention is paid to data integration (to inter-connect structured data from different sources) and data cleansing (to remove noise from database contents or discard useless records).
- *Knowledge processing services*, which exploit most prominent large data processing technologies to offer functionalities such as high performance literature mining, targeted data indexing, mining (classification, clustering, subgroup discovery etc.), abstraction and interpretation. Advanced text mining techniques such as relation extraction, similarity learning and opinion mining help to extract valuable semantic information from unstructured texts.
- *Collaborative research services*, which facilitate and augment the synchronous and asynchronous collaboration of researchers and stakeholders through adaptive workspaces, efficiently handle the representation and visualization of the outcomes of the abovementioned services through alternative and dedicated data visualization schemas, enable appropriate knowledge co-production and gaining of insights, and accommodate the orchestration of a series of actions for the appropriate handling of diverse research issues. These services enhance (both individual and group) sense-making and decision making by supporting stakeholders in: (i) expressing, arguing about, and meaningfully interacting with

relevant information and knowledge, (ii) handling data provenance and trust issues, and (iii) monitoring and comprehending the evolution of collaboration.

## 4 Advancing the State-of-the-Art in Collaborative Healthcare Research

CoHeRe is able to advance the state-of-the-art in healthcare research environments across the following three dimensions:

- *Technological dimension:* It concerns the large-scale integration and interoperability of diverse healthcare research resources. Adopting open standards and semantic web technologies, CoHeRe guarantees their conceptual and technical integration, while properly addressing a variety of issues including openness, flexibility, provenance, and workflow management.
- *Collaborative dimension:* Exploiting and building on the synergy between human and machine reasoning capabilities, CoHeRe provides innovative services to augment the efficiency, effectiveness and overall quality of collaborative research activities, thus enabling healthcare stakeholders to tame sense making, situational awareness, argumentation, knowledge co-production and collective decision making issues.
- *Organizational dimension:* Adopting suggestions and directives manifested in recent EU studies and initiatives, such as eResearch2020 (<http://www.eresearch2020.eu>) and GRDI2020 (<http://www.grdi2020.eu>), CoHeRe thoroughly addresses the issues of sustainability, governance and adoption of the foreseen working environment.

### 4.1 Technological Dimension

According to a recent study [5], scientific collaborations become “increasingly global, multipolar and networked”, something that calls for “innovative, dynamic and ubiquitous research supporting environments where scattered scientists can seamlessly access data, software, and processing resources managed by diverse systems in separate administration domains through their web browser”. A variety of systems have been already developed towards this trend, often exploiting grid computing and cloud computing technologies, ranging from ad hoc portals with limited access to content resources to general-purpose management systems with advanced services defined over a wide range of resources [6]. While useful in addressing particular needs of diverse research communities, current systems exhibit a series of limitations with respect to the functionalities offered:

- First of all, the vast majority of them miss the dynamicity and openness features of a virtual research environment, in that they should not only provide an environment for hosting, indexing, and retrieving large data sets but also leverage Web 2.0 technologies and social networking solutions to give researchers an all-inclusive environment for teamwork and resource discovery [7].
- Second, they do not provide the appropriate framework to fully support the associated “communities of practice”, which are groups of researchers and stakeholders who have similar interests or goals, and are willing to share their knowledge, insights and experiences about specific aspects that have to be collaboratively addressed [8]. These groups are characterized by informal structures and volatile memberships, an attitude of sharing information and experiences, and follow the “social theory of learning” paradigm (learning as social participation).
- Third, current systems fail to offer a unified and virtualized view of resources that come from different providers. In most cases, their scope is strictly constrained, rendering them inflexible to accommodate evolving needs and associated resources. Thus, they are unable to promote and maximize sharing and reuse of related resources.
- Fourth, they do not have the required control over intermediate and final research outcomes. They pay no or very limited attention to the issues of research results’ ownership, provenance and attribution, thus averting stakeholders’ interest in contributing to collaborative research efforts [9].
- Fifth, although there is a large number of scientific workflow management systems, such as *Taverna* (<http://www.taverna.org.uk/>) or *Kepler* (<https://kepler-project.org/>), there is a lack of open and standard-based workflow systems applied in healthcare research. Workflow solutions should be adapted to each domain, since workflow requirements in science are highly heterogeneous compared to business workflows. Specific tools are needed to facilitate data interoperability, standard-based compatibility, the right level of abstraction for the domain, etc.
- Sixth, current integration approaches are not capable of handling all possible cases of heterogeneity. Database annotations are designed to cover a very specific type of semantic equivalences, based on the classical element-to-element approach inherited from relational models. RDF however allows a more complex modeling of data and requires more advanced approaches for their semantic homogenization [10]. Similarly, low usability of interfaces for building queries for data integration systems hinders the deployment in real world environments. Advanced solutions must be found to foster their acceptance among the research community.
- Finally, current systems do not properly deal with the full spectrum of issues concerning interoperability of resources, which span from organizational to semantic and technological ones. The semantic underpinning of the underlying resources does not shape a unified space of data sets, services and stakeholders.

CoHeRe provides contemporary research communities with an innovative web-based working environment that abolishes the above limitations. It serves the different resource integration requirements through alternative integration types. It offers a highly flexible and customizable solution, enabling any researcher to easily define his/her own workflow and accordingly assemble his/her working environment with the corresponding resources without worrying about technical details. Adopting a generic approach, based on a unified and virtualized view of resources that may come from different providers, it is able to serve the highly evolving needs of healthcare researchers.

CoHeRe promotes and maximizes sharing and reuse of existing resources towards serving the diverse needs of a research community. In addition, CoHeRe provides a set of innovative resource management services, enabling researchers to locate the data and services they need through system-generated recommendations, as well as innovative competence management services, aiming to sustain the involvement of appropriate stakeholders in successful collaborative research activities.

As regards scientific workflow management, CoHeRe will collect requirements to capture the analytical steps describing computational experiments in healthcare research. Different alternatives will be analyzed among scientific and business workflow management systems, and additional tools will be developed based on widely adopted standards. Also, the possibility of generating automatic data-driven links among different tools by using domain specific vocabularies and ontologies will be studied. A portal-based tool will provide access to specific healthcare workflows that will allow researchers to easily automate large-scale processes, including data management, analysis, simulation and visualization.

As far as semantics underpinning is concerned, CoHeRe elaborates the role of semantic tags to guarantee the conceptual integration of associated data, models, services, and stakeholders. Exploiting sustainable semantic web technologies and Linked Open Data initiatives, CoHeRe meaningfully addresses the diversity of resource interoperability issues (at the data, knowledge and collaboration layers). Through this unified view of research resources, CoHeRe also offers a sustainable solution to the issues of data and research provenance and attribution, by developing and appropriately deploying novel mechanisms for handling provenance metadata.

## 4.2 *Collaborative Dimension*

The term “collaboration support software” refers to software that is designed to support a group of people involved in a common task to achieve their goals. The emergence of the Web 2.0 (and associated Social Media) era introduced a plethora of collaboration tools, which enable engagement at a massive scale and feature novel paradigms. These tools cover a broad spectrum of needs ranging from knowledge exchanging, sharing and tagging, to social networking, group authoring,

mind mapping and discussing. For instance, *Delicious* (<http://delicious.com>) and *CiteULike* (<http://www.citeulike.com>) provide services for storing, sharing and discovering of user generated Web bookmarks and academic publications, respectively. A different set of applications focuses on building online communities of people who share interests and activities (social networking applications). *Google +* (<https://plus.google.com>) and *LinkedIn* (<http://www.linkedin.com>) are representative examples of this category, giving its members the ability to create profiles which capture their interests. Special services take these interests into consideration and can suggest users with similar interests, thus facilitating the formation of online communities. Another set of Web 2.0 tools aims to collectively organize, visualize and structure concepts via maps to aid brainstorming and problem solving. Their particular emphasis is on the visual appearance of the workspace. Tools such as *Thinkature* (<http://www.thinkature.com>) and *FreeMind* (<http://freemind.sourceforge.net>) fall into this category.

Although all the above tools enable the massive and unconstraint collaboration of users, this very feature is the source of a problem that these tools introduce: the problem of information overload. The amount of information produced and exchanged and the number of events generated within these tools exceeds by far the mental abilities of users to: (i) keep pace with the evolution of the collaboration in which they engage, and (ii) keep track of the outcome of past sessions. As the same user may be using different tools for different collaboration sessions, e.g. social networking tools in one situation and discussion forums in another, the problem of information overload stretches across such “channels” amplifying its impact. Current Web 2.0 collaboration tools exhibit two important shortcomings making them prone to the problems of information overload and cognitive complexity. First, these tools are “information islands”, thus providing only limited support for interoperation, integration and synergy with third party tools. While some provide specialized APIs with which integration can be achieved, these are primarily aimed at developers and not end users. Second, Web 2.0 collaboration tools are rather passive media, in the sense that lack reasoning services with which they could actively and meaningfully support collaboration. Support for reasoning or related mechanisms could facilitate problem solving and decision making in these tools, especially since they do not appropriately scale to deal with voluminous and complex data.

As far as argumentative collaboration is concerned, various tools focusing on the sharing and exchange of arguments, diverse knowledge representation issues and visualization of argumentation have been developed. Tools such as *Compendium* (<http://compendium.open.ac.uk>) allow users to create issues, take positions on these issues, and make pro and contra arguments. They can capture the key issues and ideas and create shared understanding in a knowledge team; in some cases, they can be used to gather a semantic group memory. However, current tools supporting argumentative collaboration have the same problems with the aforementioned Web 2.0 collaboration tools. They too are considered standalone applications thus being “information islands”, lacking support for interoperability and integration with other tools (e.g. with data mining services foraging the Web to discover interesting

patterns or trends). They also cope poorly with voluminous and complex data as they provide only primitive reasoning services. This makes these tools prone to the problem of information overload. Moreover, argumentative collaboration support tools reveal additional shortcomings that prevent them from reaching a wider audience. In particular, their emphasis on providing fixed and prescribed ways of interaction within collaboration spaces makes them difficult to use as they constrain the expressiveness of users. This resulted in making these systems being used only in niche communities [11].

CoHeRe may fully cover the diversity of collaborative research requirements in contemporary settings by providing a series of innovative features. Firstly, it provides advanced collaboration support functionalities through innovative virtual workspaces offering alternative visualization schemas. Such visualizations will facilitate the processes of perceiving, monitoring and comprehending the evolution of collaboration, offering augmented awareness about the issue at hand. Moreover, alternative visualizations will help stakeholders control the impact of voluminous and complex data. Secondly, CoHeRe does not treat collaboration services as standalone applications that operate autonomously, but rather as ones that coexist and make use of other services in order to improve their performance. For instance, CoHeRe collaboration services will be able to meaningfully accommodate the outcomes of literature mining services in a collaboration session, thus supporting evidence-based collaboration. Thirdly, by acknowledging that collaborative research issues have to be solved through dialoguing and argumentation among stakeholders, CoHeRe provides a novel argumentative collaboration support service geared towards achieving consensus and gaining of insights. Fourthly, by supporting emergent semantics and the incremental formalization of argumentative collaboration, CoHeRe exploits the synergy between human and machine reasoning to facilitate and augment individual and collective sense-making and decision making. By providing ease-of-use and expressiveness for users and advanced reasoning by the machine, CoHeRe integrates appropriate recommendation mechanisms that enable stakeholders to engage the appropriate resources in their dynamic work settings. Overall, CoHeRe offers an innovative collaborative environment that allows users to “immerse” in Web 2.0 interaction paradigms and exploit its enormous potential to collaborate through reviewing, commenting on and extending the shared content. The proposed environment enables stakeholders to maintain chains of views and opinions (accompanied by the supporting data), which may reflect at any time the current collective knowledge on the issue under consideration, and justify a particular decision made or action taken.

### ***4.3 Organizational Dimension***

As emphasized throughout recent scientific reports of the GRDI2020 initiative, sustainability and adoption are two major challenges affecting Virtual Research Environments (VREs) development. With respect to sustainability, it is broadly

admitted that the majority of current solutions does not have a long term support and only serves needs of specific communities [1]. They are not open enough to become an integral part of long term research efforts, as communication networks and high-performance computing are. Moreover, they often integrate services that are highly specific (i.e. useful for a particular community to address a particular problem), rendering them not generic enough to solve common problems and thus be embraced by communities across disciplines and institutions.

As far as adoption is concerned, the above reports underline that the majority of current VREs are not yet fully integrated into standard practices, tools and research protocols used by real life communities of practice. This is due to various reasons including inadequate user support, gaps between the real research community needs and the actual services integrated in them, reliability of technology used, poor user engagement, legal and ethical issues, and diversity in ways of working and “languages” spoken across disciplines.

Acknowledging that the development and management of VREs is not just a technological process but also a societal and organizational one, CoHeRe will thoroughly address the above challenges. It will develop a clear business model to sustain the foreseen solution relying on both internal and external funding and support. It will provide structures and shape policies for the governance of the CoHeRe working environment to enable its efficient and effective use by research communities across institutions and borders. It will be designed for reuse, exploiting common building blocks and relying on existing infrastructures as much as possible, while also being flexible to allow the user-friendly integration of components for specific use.

The management of the overall CoHeRe working environment will not require any particular expertise. Novel services for resources management, offering among others recommendations about the targeted exploitation of resources and researchers' competences, will be developed. Much attention will be paid to the adoption of open standards and custom web technologies to ensure reliability. A range of innovative functionalities for collaborative research support will be offered to augment user engagement and accommodate interdisciplinary ways of working.

Finally, with respect to legal and ethical issues (mainly for patients data from related clinical studies), CoHeRe accommodates and appropriately adapts a Data Protection Framework that relies primarily on pseudonymization and de facto anonymization of personal data. A strong security framework, binding contracts, the use of a Trusted Third Party and the procurement of appropriate informed consent ensure that the pseudonymized data cannot be linked back to the original patient with reasonable means and can thus be considered de facto anonymous.

## 5 Conclusion

By advancing the state-of-the-art in virtual research environments, CoHeRe has the potential to become a sustainable healthcare research ecosystem that acts as both a consumer and a provider of research resources. It provides a novel collaborative environment that is flexible enough to meaningfully accommodate heterogeneous tasks, such as data analytics, in silico modeling, knowledge discovery and collective sense making, which are typically handled by separate systems. This alleviates expenses related to the large-scale data loading into multiple systems and eases the management of big data volumes, which are among the basic concerns and first class priorities of scientific research communities. Due to the open and flexible development approach followed in CoHeRe, which takes into account the way science is organised and run, the proposed solution enables a digital science ecosystem for both individual disciplines and for multidisciplinary work by supporting individual, dynamic communities of research, with their associated tools, networks and practices.

## References

1. Carusi, A., Reimer, T.: Virtual Research Environment Collaborative Landscape Study, JISC (2010)
2. Karacapilidis, N.: Integrating New Information and Communication Technologies in a Group Decision Support System. *Int. Trans. Oper. Res.* **7**(6), 487–507 (2000)
3. Karacapilidis, N., Tzagarakis, M., Karousos, N., Gkotsis, G., Kallistros, V., Christodoulou, S., Mettouris, C., Nousia, D.: Tackling cognitively-complex collaboration with CoPe\_it! *Int. J. Web-Based Learn. Teach. Technol.* **4**(3), 22–38 (2009)
4. Karacapilidis, N. (ed.): Mastering Data-Intensive Collaboration and Decision Making: Cutting-edge Research and Practical Applications in the Dicode Project. Studies in Big Data Series, vol. 5. Springer (2014)
5. Llewellyn Smith, C., Borysiewicz, L., Casselton, L., Conway, G., Hassan, M., Leach, M., et al.: Knowledge, Networks and Nations: Global Scientific Collaboration in the 21st Century. The Royal Society (2011)
6. Candela, L.: Virtual Research Environments. GRDI2020 Scientific Report (2011)
7. Allan, R.: Virtual Research Environments: From Portals to Science Gateways. Chandos Publishing, Oxford (2009)
8. Wenger, E., Snyder, W.: Communities of practice: the organizational frontier. *Harvard Bus. Rev.* **78**, 139–145 (2000)
9. De Roure, D., Goble, C., Stevens, R.: The design and realisation of the my experiment virtual research environment for social sharing of workflows. *Fut. Gen. Comput. Syst.* **25**, 561–567 (2009)
10. Anguita, A., García-Remesal, M., de la Iglesia, D., Graf, N., Maojo, V.: Toward a view-oriented approach for aligning RDF-based biomedical repositories. *Methods Inf. Med.* **54**(1), 50–55 (2015)
11. Karacapilidis, N., Koukouras, D.: A web-based system for supporting collaboration towards resolving oncology issues. *Oncol. Rep.* **15**(4), 1101–1107 (2006)

# An Ontology-Based Approach for Representing Medical Recommendations in mHealth Applications

Aniello Minutolo, Massimo Esposito and Giuseppe De Pietro

**Abstract** Nowadays, mHealth applications have been evolving in the form of pervasive solutions for supporting healthy life-style and wellness self-management. In such a direction, the Italian project “Smart Health 2.0” realized innovative technological infrastructures, on which different mHealth applications and services were developed, aimed at remotely supporting individuals in diseases prevention and improving their welfare and life styles. In this paper, the ontology-based approach proposed in the project to represent, share, and reason on the knowledge characterizing a subject within mHealth applications is presented. The proposed approach uses a hybrid strategy integrating ontology models and deductive rules built on the top of them. In order to better describe the proposed approach, a case of application has been presented with respect to an mHealth application designed for managing diet according to given daily caloric needs.

**Keywords** Ontology • Logic rules • mHealth • Knowledge-based systems

## 1 Introduction

Nowadays, the growing penetration of mobile devices, coupled with infrastructures for telecommunication, has deeply influenced the delivery of healthcare services [1], and defined wider horizons for health through mobile technologies [2]. The cheap and widespread availability of mobile phones and wearable devices has enabled the development of new mobile health (mHealth) systems, deployed on

---

A. Minutolo (✉) · M. Esposito · G. De Pietro

Institute for High Performance Computing and Networking, ICAR-CNR, via P. Castellino,  
111-80131 Naples, Italy  
e-mail: minutolo.a@na.icar.cnr.it

M. Esposito  
e-mail: esposito.m@na.icar.cnr.it

G. De Pietro  
e-mail: depietro.g@na.icar.cnr.it

mobile phones provided to the individuals, able to deliver innovative health services for improving the individual's comfort, enhancing the quality of life [3], promoting wellness and lifestyles [4], or improving the adherence to therapies of remote monitored patients [5].

The Italian project “Smart Health 2.0” operated in such a direction, aimed at the realization of innovative technological infrastructures, on which to develop different services and facilities dedicated to the implementation of new business models in the area of health and well-being. In detail, a remote and mobile framework was developed where several mHealth applications and services communicated and cooperated with the goal of remotely supporting individuals in diseases prevention and improving their welfare and life styles. mHealth applications developed in the project were aimed at supporting the individuals in the prevention and/or management of Diabetes Mellitus Type 2, Migraine frequent attacks, dysphonia, and wellness.

Each mHealth application is dedicated to the management of a specific health area of interest, and has to be deployed on the user smartphone for locally monitoring his/her activities. The data collected and processed by each mHealth application is, on the one hand, locally evaluated in order to support users in the monitoring of their activities and to provide them tailored recommendations about their lifestyles with respect to the estimation of the risk to contract a disease or to experience a worsening of their healthy status wellness. On the other hand, the information evaluated by each application can be shared with a centralized server where data are integrated and fartherly evaluated in order to compose more complex and structured data, and persistently store them.

This complex scenario demands for an appropriate knowledge model for properly representing, sharing, and reasoning on the knowledge characterizing the individual context and, thus, two main issues arise: (i) how to handle and exchange the processed knowledge among existing mobile and remote knowledge sources; (ii) how to evaluate and reason on the acquired knowledge directly on mobile devices for locally performing an accurate and continuous analysis of the individual’s health status.

In this paper, the ontology-based approach proposed to address these issues in such a project is presented. In particular, a knowledge model is produced, which is applicable in the domain of every involved mHealth application. In fact, it enables the easy comparison and integration of information gathered from different measurements and/or applications [6]. Moreover, since not all information provided by a mHealth application is relevant to other ones, it is able to provide both a common semantics for representing knowledge among heterogeneous actors, and a common and upper layer on top of which each actor can formulate its unique and domain-specific knowledge [6].

Finally, it enables inference mechanisms on top of the formalized knowledge with the goal of deriving new and higher-level knowledge from the acquired low-level one. In fact, the information acquired from users and/or devices such as sensors constitutes a low-level knowledge and, in the case when inference mechanisms are locally applied inside a single mHealth application, it can be evaluated

for obtaining higher-level knowledge and better decisions for supporting users and their activities.

In the following, Sect. 2 provides some preliminary information and details why an ontology-based approach has been chosen for the project. In Sect. 3, the proposed ontology-based approach is presented, whereas its application to a case of study is described in Sect. 4. Finally, Sect. 5 concludes the work.

## 2 Background and Motivations

Knowledge pertaining the domain of every mHealth application involved in the project could be modeled via different techniques [7–9], starting from key-value models, where relevant knowledge is simply represented as a list of attributes and their assigned values, to markup scheme models where complex hierarchical data structures are composed via dedicated languages such as XML and/or RDF [10]. Such techniques provide basic reasoning methods, but they still possess some limitations with respect to their inferring ability and their offered semantic and expressive capacity [8].

The employment of ontology models is one of the most used approaches for representing the knowledge handled by medical applications, especially in the case of mobile scenarios, due to their capacity to provide both formal and semantically rich mechanisms for describing complex data structure and for enabling logic and deductive reasoning facilities upon on them [8, 9, 11]. Ontology models use simple description logic languages to represent domain-specific concepts in terms of characterizing attributes and semantic relationships existing among them. For this reason, they were identified as adequate to meet the requirements of the project in terms of expressivity, formality, reusability, and inference abilities. In fact, they let to express knowledge in an expressive and formal language that facilitates the sharing of information among heterogeneous actors.

In more detail, domain-specific knowledge of each application can be distilled in terms of concepts and properties extending the predefined formal basis. Moreover, they offer the chance to compose upper ontologies where most high-level and common knowledge is modeled and, on top of which, distinct applications can create their unique and domain-specific ontologies where the relevant information pertaining specific applications is distilled. Finally, on top of ontology models, logic and deductive rules can be formalized in order to provide logical reasoning mechanisms. This aspect is fundamental since new and high-order knowledge can be inferred by evaluating low-order knowledge formalized in an ontology model, such as the evaluation of context by checking for consistency, compatibility, incompleteness and ambiguity [9, 12].

### 3 The Proposed Approach

The approach here proposed uses a hybrid strategy where both ontology models and logic deductive rules are involved, for formally representing the knowledge characterizing medical recommendations. Three main steps constitute the approach: (i) definition of an upper ontology model for formalizing the common knowledge characterizing the scenario and the actors involved in it; (ii) specialization and enrichment of the upper ontology model for representing the specific knowledge characterizing medical recommendations related to an mHealth application; (iii) definition of logic rules operating on top of the composed ontology model, in order to represent the procedural knowledge underlying the medical recommendations to reproduce.

In detail, for each mHealth application, a dedicated ontology model is composed by integrating the upper ontology model with a specialized one defined for formally describing the relevant concepts involved in the medical recommendation to formalize. On the one hand, the ontology model represents such concepts, their attributes of interest, and the semantic relations existing among them. On the other hand, rules describe the deductive logic underling the medical recommendations, i.e. the logic used for determining when and how they have to be generated taking into account the current context of user and his/her health status. The use of an upper and common ontology model, where the common knowledge characterizing the scenario of application is formalized, enables, on the one hand, the easy exchange of information among heterogeneous actors of the scenario, and, on the other hand, the easy integration of knowledge for composing enriched functionalities build on top of knowledge coming from distinct sources.

As discussed in the previous section, ontology-based approaches provide a formal and standard framework for describing the knowledge characterizing a complex scenario and the context of its actors. Nevertheless, ontology models can become computationally expensive for complex domains [8]. With respect to this issue, the proposed approach let to reduce the number of classes and properties processed by each mHealth application, since just the required portion of the whole knowledge base is managed and evaluated. Moreover, since ontology models typically demand extensive computational effort to maintain decidability, the N-Triples serialization of OWL language [13] has been used in the project. This solution, on the one hand, ensures decidability over expressive power since OWL is a widely used subset of description logics, and, on the other hand, the N-Triples serialization offers a light-weight format for ontologies, which results easier to parse and process in mobile applications [14].

Finally, with respect to the definition of logic rules on top of ontology models, the proposed approach makes use of a subset of the Jena rule language [15], since it is suitable for being efficiently parsed on resource-limited settings, and its predicates can be selected among classes and properties encoded as N-Triples serialization of OWL.

## 4 The Application to the Smart Health 2.0 Scenario

The mHealth applications developed in the project aim at evaluating the health status of monitored users by collecting and processing sets of measures, which can be automatically acquired via sensors or manually inserted via assisted questionnaires.

According to the proposed approach, each mHealth application has been designed as a rule-based application where the knowledge underlying its functionalities is formalized via ontology models with the goal of easily sharing and interpreting the knowledge produced. Thus, as a first step of the proposed approach, the application scenario has been analyzed in order to determine the common knowledge characterizing it so to formalize such knowledge inside an upper ontology model.

### 4.1 *Definition of the Upper Ontology Model*

The health status of a subject can be characterized both by time-independent, or slowly variable, knowledge, such as age or daily caloric needs, and by time-dependent knowledge describing acquired measures whose values are characterized by the time of acquisition. Such measures are typically compared with the medical recommendations that are selected as relevant for the user according to his/her context. For instance, given a user having  $x$  calories as daily caloric needs, just the recommendations dedicated for  $x$  calories have to be selected as relevant for him/her.

Thus, when his/her health status is evaluated, as first step, the set of modeled recommendations is evaluated in order to assign the pertaining ones to the user. After, the health status is evaluated by analyzing the recommendations selected together with slowly variable attributes and acquired measures in order to recognize potential abnormal situations, suggestions to provide, or goals to achieve. Abnormal situations and suggestions are explicit interactions with the mHealth apps since they have to be notified in order to produce an alarm or visualize a dialog for showing the outcomes of the health evaluation. Goals, instead, are time-dependent knowledge to be added to the health status of the user for applying time-based recommendations, i.e., for instance, it is recommended to repeat an evaluation after a period of time is elapsed in order to determine if a worsening is happened.

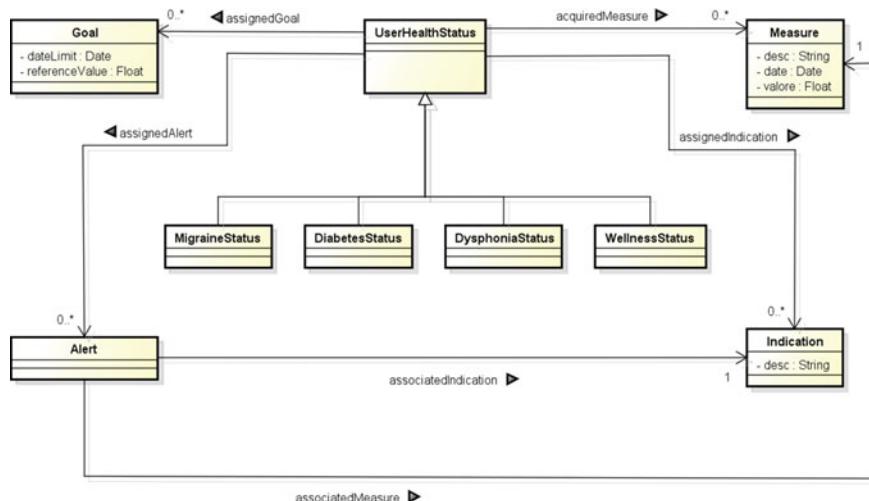
Starting from these considerations, an upper ontology has been defined for representing the following concepts:

- **UserHealthStatus:** it models the current health status of a subject in terms of information to consider or generated during the evaluation of his/her health conditions.
- **Indication:** it models an indication that has to be applied during the current evaluation. At this level, it is just a high-level concept containing a text

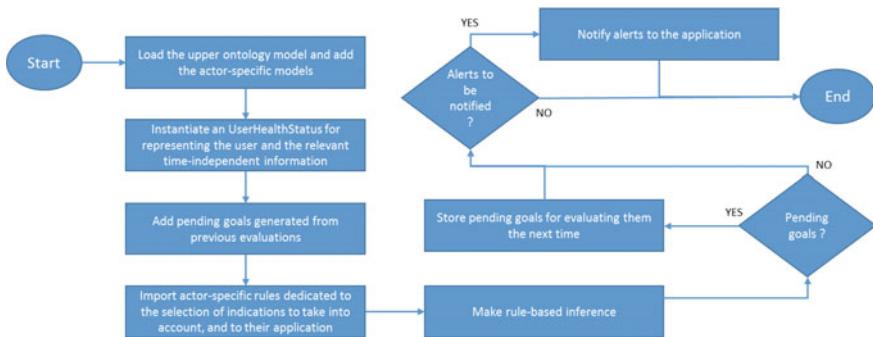
describing the indication. At a lower level, it is specialized for modelling information closer to measures to evaluate and to the logic to reproduce.

- **Measure:** it models a measure collected from the app that is of interest for the current health evaluation, and it is characterized by the time of acquisition.
- **Goal:** it models a goal to achieve in a given period of time. A goal is characterized by the time it must be verified by, and, eventually, by a reference value to determine when the goal is met.
- **Alert:** it models an alert to be notified to the user. Generally, an alert is characterized by a measure, and by the indication that justifies the alert itself.

Figure 1 reports the main concepts defined for composing the upper ontology model used for representing the common knowledge that each actor is able to interpret and integrate with the knowledge that it produce. This upper ontology is used by every actor of the scenario that is in charge of specializing the upper ontology model and reasoning on top of it in accordance with the upper-level reasoning scheme reported in Fig. 2. In detail, each mHealth application, on the one hand, enriches the upper ontology model specializing the concepts characterizing the health domain of its interest, and, on the other hand, it introduces a set of production rules able to interpret and operate on the composed ontology models. In this respect, in order to better describe how the knowledge characterizing an mHealth specialization has been formalized by integrating and specializing the defined upper ontology model, in the next section a closer look to the mHealth application designed for managing the diet will be presented and discussed.



**Fig. 1** Upper ontology model for representing common knowledge of the scenario considered



**Fig. 2** The upper-level reasoning scheme applied on the *top* of the upper ontology model

#### 4.2 Specialization and Enrichment of the Upper Ontology Model

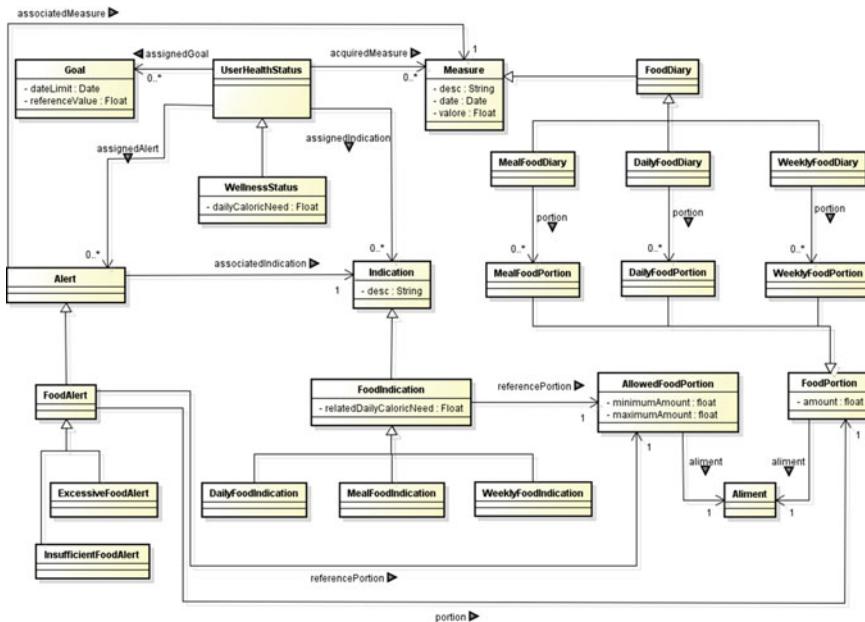
In order to support users in keeping their wellness and health lifestyles, an mHealth application has been designed for taking their diet under control according to given daily caloric needs. In detail, the mHealth application supports users to monitor the food portions consumed during a meal, over a week of observation. For each aliment, a set of diet recommendations has been formulated describing the allowed portion for a meal, the allowed portion for a day, and the allowed portion for a week. Allowed portions are distilled in terms of minimum and maximum portions that are recommended and, respectively, forbidden to consume in a given period of time.

The goal of the mHealth application is to recognize and alert to the user potential abnormal situations associated to the consumed food portions, such as the inadequate and excessive consumption of aliments, with respect to the caloric need of the user. In accordance with the proposed approach, this mHealth application has been designed as a rule-based application where the knowledge is formally represented and encoded via ontology models and logic rules operating on top of them.

In detail, starting from the upper ontology model defined in the previous section, the mHealth ontology model has been composed by integrating the specific knowledge underlying the diet recommendations to formalize. In this respect, Fig. 3 reports the main concepts defined for composing the knowledge required for formulating the diet guidelines underlying the considered mHealth application.

With more details, the main concepts defined for composing the mHealth ontology model can be described as follow:

- **WellnessStatus:** it models the current status of wellness of a subject. It is represented as a specialization of the concept UserHealthStatus having further information pertaining the user, such as his/her daily caloric need.
- **FoodPortion:** it models a food portion about a single aliment that is consumed, and also the grams consumed. The FoodPortion concept is further specialized



**Fig. 3** The mHealth ontology model for representing its domain-specific knowledge

for indicating the time in which the food has been consumed, i.e. in a single meal (**MealPortion**), in a single day (**DailyPortion**), in an entire week (**WeeklyPortion**).

- **Aliment:** it models the aliments allowed to consume and that can be associated to the food portions consumed by the user.
- **FoodDiary:** it models a diary where the set of food portions consumed during a given time is annotated. It is worth noting that FoodDiary is defined as a specialization of Measure since it models a time-dependent observation about the monitored user. This concept has been further specialized in **Meal-FoodDiary** for representing a food diary about the portions consumed during the same meal; **DailyFoodDiary** for representing a food diary about the portions consumed in a single day, and it contains both the single meals (**Meal-FoodDiary**) and the total amount of portions (**DailyPortion**) consumed during the day; **WeeklyFoodDiary** for representing a food diary about the portions consumed in an entire week, and it contains both the daily diaries (**Daily-FoodDiary**) and the total amount of portions (**WeeklyPortion**) consumed in a week.
- **AllowedPortion:** it models a food constraint about a given aliment, represented as minimum and maximum amounts in grams to consume.
- **FoodIndication:** it models a food recommendation about the portion of a given aliment that is suggested to consume during a single meal, during a single day, or during an entire week. It is worth noting that FoodIndication is a

specialization of Indication defined in the upper ontology model. Each FoodIndication is characterized by a daily caloric need, and is associated to an AllowedPortion expressing both the aliment to which the food indication is related, and the allowed amount to consume. FoodIndication is further specialized in order to model food indications pertaining a meal portion (**MealFoodIndication**), a daily portion (**DailyFoodIndication**), or a weekly portion (**WeeklyFoodIndication**).

- **FoodAlert**: it models an alert to be notified to the user when an abnormal consume of food has been recognized. Food alerts are generated due to a given food indication, and contain also the potential abnormal portion and the associated allowed portion violated. A FoodAlert is further specialized in **ExcessiveFoodAlert** and **InsufficientFoodAlert** to model excessive and insufficient food consumptions in a given portion, respectively.

### 4.3 Definition of Logic Rules Operating on Top of the Composed Ontology Models

The goal of this step is to formulate a set of logic rules on top of the defined mHealth ontology model able to realize the procedural knowledge underpinning the diet recommendation to reproduce. The logic rules are distilled as if-then rules operating on the ontology models, and are expressed via a light-weight as well as expressive formalism, whose grammar is reported in Fig. 4.

A rule is made by a conjunction of condition elements in its left-hand side (LHS), and a set of actions in its right-hand side (RHS). Each rule has to enable: (i) the evaluation of the health status of the user and his/her slowly variable knowledge in order to determine the most pertaining recommendations; (ii) the evaluation of the selected recommendations and his/her time-dependent information, i.e. the collected measures, in order to detect potentially abnormal measures; (iii) the generation of alerts in accordance with the recognized abnormal measures.

```

Rule ::= Name: if Antecedents then Consequents;
Name ::= TextValue;
Antecedents ::= ConditionElement [conjunctionOperator ConditionElement];
Consequents ::= ActionElement [conjunctionOperator ActionElement];
conjunctionOperator ::= , | &
ConditionElement ::= (N-Triple) | not(N-Triple) | FunctionToCall
ActionElement ::= N-Triple | NonMonotonicFunction;
N-Triple ::= Subject Predicate Object;
Subject ::= Term | Variable; Predicate ::= Term | Variable; Object ::= Term | Variable | Value;
Term ::= uri-reference (e.g. http://exuri#ex1) | prefix:name (e.g. expre:ex1);
Variable ::= ?variableName; Value ::= NumericalValue | BooleanValue | TextValue; fArg ::= Variable | Value;
NumericalValue ::= number (e.g. 10 or 20.5); BooleanValue ::= 'true' | 'false'; TextValue ::= 'a string';
FunctionToCall ::= LogicalFunction | ArithmeticalFunction | NonMonotonicFunction;
NonMonotonicFunction ::= newInstance(Variable, Term) | retract(N-Triple);
LogicalFunction ::= equal(fArg, fArg) | notEqual(fArg, fArg) | lessThan(fArg, fArg) |
                  lessEqual(fArg, fArg) | greaterThan(fArg, fArg) | greaterEqual(fArg, fArg);
ArithmeticalFunction ::= sum(fArg, fArg, Variable) | product(fArg, fArg, Variable) |
                      difference(fArg, fArg, Variable) | quotient(fArg, fArg, Variable);

```

**Fig. 4** The rule language used for encoding logic rules on *top* of ontology models

With respect to this first requirement, a rule for evaluating the user daily caloric need and selecting the proper diet recommendation can be encoded as reported below:

```
[R1: (?y rdf:type base:WellnessStatus), (?y base:dailyCaloricNeed ?c1),
  (?ind rdf:type base:FoodIndication), (?ind base:relatedDailyCaloricNeed ?c2), equal(?c1,?c2)
-> (?y base:assignedIndication ?ind), (?y base:contieneIndAlimSingoloPasto ?ind)]
```

With respect to the second requirement, a set of production rules has been formulated for evaluating the user's food diaries and the consumed food portions annotated in them. At this aim, as an example, a consumed food portion annotated in a meal food diary can be successively included inside both a daily food diary, and inside weekly food diary. Thus, as an example, potential abnormal meal food portions have to be researched in any kind of food diary associated to the user wellness status.

Finally, in order to satisfy the last requirement, logic rules must be also organized to generate the proper alert when an abnormal food portion is recognized, by discriminating over insufficient and excessive consumptions and the corresponding alerts to generate. As an example, a production rule for identifying insufficient consumptions in the food portions annotated inside a meal food diary, can be encoded as reported below:

```
[R2: (?y rdf:type base:WellnessStatus), (?y base:assignedIndication ?ind),
  (?ind rdf:type base:MealFoodIndication), (?ind base:referencePortion ?r),
  (?r base:aliment ?paAlim), (?y base:acquiredMeasure ?mfd),
  (?mfd rdf:type base:MealFoodDiary), (?mfd base:portion ?p), (?p base:aliment ?paAlim)
  (?p base:amount ?cA), (?r base:minimumAmount ?minA), lessThan(?cA,?minA)
-> newInstance(?al, base:InsufficientFoodAlert), (?y, base:assignedAlert, ?al),
  (?al base:associatedMeasure ?mfd), (?al base:portion ?p), (?newAl base:referencePortion ?r) ]
```

## 5 Conclusions

In this paper, the ontology-based approach proposed in the Italian project "Smart Health 2.0" was presented, whose aim is to represent, share, and reason on the knowledge characterizing a subject within mHealth applications.

The proposed approach uses a hybrid strategy where both ontology models and deductive rules are involved, for formally representing the knowledge characterizing a given domain. In detail, the three main steps are foreseen: (i) the definition of an upper ontology model for representing the common knowledge characterizing the scenario and the actors involved in it; (ii) the specialization and enrichment of the upper ontology model for representing the specific knowledge characterizing an

mHealth application; (iii) the definition of logic rules operating on top of the composed ontology model, in order to represent the procedural knowledge underlying the recommendations to reproduce. In order to better describe the proposed approach, a case of application has been presented with respect to an mHealth application designed for managing diet according to given daily caloric needs. Next step of the research activities will regard the evaluation of performances of mobile knowledge-based systems with the goal of designing innovative and formal optimization procedures to be integrated with the proposed approach. The final aim will be the revision of the described ontology models and rules for diminishing the portion of knowledge base evaluated in the reasoning process.

**Acknowledgments** This work has been partially supported by the Italian project “Smart Health 2.0” funded by the Italian Ministry of Education, University, and Research (MIUR).

## References

1. Malvey, D.M., Slovensky, D.J.: *mHealth: Transforming Healthcare*. Springer, (2014)
2. World Health Organization: *mHealth: New horizons for health through mobile technologies*. Global Observatory for eHealth Series, vol. 3, Geneva, Switzerland, (2011)
3. Akter, S., et al.: Modelling the impact of mHealth service quality on satisfaction, continuance and quality of life. *Behav. Inf. Technol.* **32**(12), 1225–1241 (2013)
4. Knight, E., Stuckey, M.I., Petrella, R.J.: Health promotion through primary care: enhancing self-management with activity prescription and mHealth. *Phys. Sportsmed.* **42**(3), 90–99 (2014)
5. Hamine, S., et al.: Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *J. Med. Internet Res.* **17**(2) (2015)
6. Krummenacher, R., Strang, T.: Ontology-based context modeling. In: Proceedings of the 3rd Workshop on Context Awareness for Proactive Systems, Guildford (2007)
7. Strang, T., Linnhoff-Popien, C.: Context modelling survey. In: Proceedings of the Workshop on Advanced Context Modelling, Reasoning and Management, UbiComp—The Sixth International Conference on Ubiquitous Computing, Nottingham, UK (2004)
8. Bettini, C., et al.: A survey of context modelling and reasoning techniques. *Pervasive Mobile Comput.* **6**(2), 161–180 (2010)
9. Nadoveza, D., Kiritsis, D.: Ontology-based approach for context modeling in enterprise applications. *Comput. Ind.* **65**(9), 1218–1231 (2014)
10. RDF/XML Syntax Specification (2004). <http://www.w3.org/TR/REC-rdf-syntax/>
11. Ejigu, D., Scuturici, M., Brunie, L.: An ontology-based approach to context modeling and reasoning in pervasive computing. In: Proceedings of the Fifth Annual IEEE International Conference On Pervasive Computing and Communications Workshops, PerCom Workshops’ 07, pp. 14–19 (2007)
12. Wang, X.H., Gu, T., Zhang, D.Q., Pung, H.K.: Ontology based context modeling and reasoning using OWL. In: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops, pp. 18–22 (2004)
13. Patel-Schneider, P., Hayes, P., Horrocks, I., et al.: OWL web ontology language semantics and abstract syntax. In: W3C Recommendation, vol. 10 (2004). <http://www.w3.org/TR/owl-semantics/>

14. Minutolo, A., Esposito, M., De Pietro, G.: Design and validation of a light-weight reasoning system to support remote health monitoring applications. *Eng. Appl. Artif. Intell.* **41**, 232–248 (2015)
15. Carroll, J.J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K.: Jena: implementing the semantic web recommendations. In: Proceedings of the 13th International World Wide Web conference on Alternate track, New York, USA, pp. 74–83 (2004)

# Semantic Cluster Labeling for Medical Relations

Anita Alicante, Anna Corazza, Francesco Isgrò and Stefano Silvestri

**Abstract** In the context of the extraction of the semantic contents important for the effective exploitation of the documents which are now made available by medical information systems, we consider the processing of relations connecting named entities and propose an unsupervised approach to their recognition and labeling. The approach is applied to an Italian data set of medical reports, and interesting results are presented and discussed from a qualitative point of view.

## 1 Introduction

Plenty of digital documents have been made available since the introduction of medical information systems. To allow a full exploitation of this material, tools are required for browsing and searching to satisfy information needs on the basis of their semantic contents. While part of the material, mainly including international scientific publications, is in English, increasingly more material is being created in the language of the country of the medical institution. The main part of the local language material is represented by patient records. They contain information important not only for preparing care plans or solve problems for the particular patient, but also to extract statistics useful for research and also for logistics administration.

---

A. Alicante (✉) · A. Corazza · F. Isgrò  
Dip. di Ingegneria Elettrica e delle Tecnologie dell'Informazione (DIETI),  
Università degli Studi di Napoli Federico II, Napoli, Italy  
e-mail: anita.alicante@unina.it

A. Corazza  
e-mail: anna.corazza@unina.it

F. Isgrò  
e-mail: francesco.isgro@unina.it

S. Silvestri  
Dip. di Medicina Sperimentale, Seconda Università degli Studi di Napoli,  
Caserta, Italy  
e-mail: stefano.silvestri@gmail.com

When accessing patient records, it is necessary to be careful about privacy issues: also in this case, a semantic interpretation of the text is crucial to elaborate effective strategies to protect patient privacy. In conclusion, semantic-driven access to such texts would allow more effective extraction and usage of the information they contain, while respecting all the constraints imposed by the nature of the documents. As discussed in [1], semantic processing of texts is usually based on recognition of named entities and relations connecting them. In the cited work, we extensively discuss a domain entity and relation recognition system for Italian. In the present work, only the entity recognition step is taken from that work, as here we only focus on the relation recognition step.

Furthermore, while considering relations, a slightly different problem is tackled with respect to the one analyzed in [1], and usually in the literature, where the type of the relation is decided for a particular occurrence of an entity pair. Here, on the other hand, rather than considering an occurrence of an entity pair in a particular context, only the two entities are taken into account to decide whether a relation connects them and how such relation can be labeled. In fact, this is complementary with respect to the task of deciding whether two entities are related, which should be decided on the basis of the context where the two entities occur, as in [1]. On the other hand, by considering only the two involved entities, we can only decide the *type* of a relation. Further analyses also considering the context in which the entities occur, could then decide whether the relation is stated or negated.

We propose to use Word Embeddings (WEs) [2] to represent the words involved in each entity with a real valued array. WEs most interesting characteristic consists in the fact that the mutual position of words in a metric space strongly depends on their meanings, so that words having similar semantics have large similarity, when this is computed, for example, by cosine similarity. For each entity, we then consider the embeddings corresponding to each token. Given the geometrical characteristics of WEs, it makes sense to compute for each entity the average of the WEs and to take the resulting vector as its representative. Furthermore, embeddings can be built from large collection of unannotated set with a very efficient algorithm. Therefore, they can be easily applied to a new language, in our case to Italian, provided that enough texts are available. We used documents extracted from Wikipedia for this.

Each pair of entities occurring in the same sentence represent a possible candidate for a relation. We therefore build the feature vector for each entity pair by juxtaposing the so-obtained average vectors for each entity and input this representation into a  $k$ -means clustering [3].

The paper is organized with Sect. 2 devoted to the discussion of related work; Sect. 3 detailing the approach implementation; Sect. 4 presents the qualitative results obtained on a data set; and Sect. 5 for conclusions and future work.

## 2 Related Works

Detecting relations connecting entities in a text has attracted attention because this task is crucial when reconstructing the text's content. Sometimes, specific kinds of relation are considered, as for example in the BioCreative V context Chemical Disease Relation Track [4], where attention is focused on the relation between a chemical and a disease it induces. In several applications, however, the possible kinds of relations are not known a-priori, but should be reconstructed together with the occurrence of a relation: see for example [1] and references therein.

Effectiveness of WEs [2] to represent semantic properties of words in a geometrical space have been advantageously exploited in several tasks in the natural language processing area. As an example in the bio-medical domain, in the BioCreative V context Chemical Disease Relation Track cited above, [5] includes word embeddings among the features used to represent the input of a conditional random field classifier for named entity recognition.

Although different approaches have been proposed for the construction of such representation, the most popular is probably *word2vec*.<sup>1</sup> This approach only considers the ability of the representation of predicting words, and it is completely independent from the considered task, that is relation recognition and classification. A diametrically different strategy has been adopted, for example, by [6], where WEs are specifically constructed for the relation classification task. In fact, we aim at understanding whether we can mine relations from texts by only using a general purpose tool like *word2vec*.

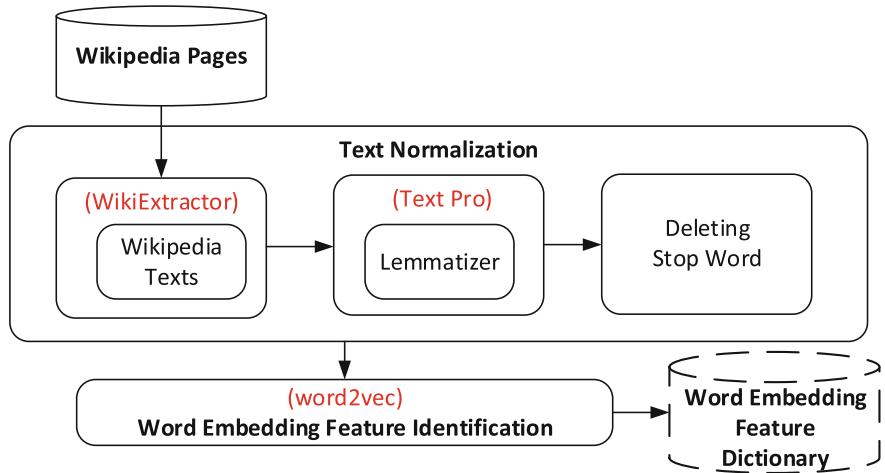
As discussed in the introduction, the task we are considering here is different from the ones in the references above, generally referred to as relation recognition and/or classification, in an important point: rather than considering a particular occurrence of the entity pair, we only consider the tokens corresponding to the two entities. This means that we are not trying to decide whether there is a relation in correspondence of a specific entity pair instance, but whether there can be a relation between two entities which can then occur in different points of the text.

## 3 System Architecture

The system proposed in this paper is structured in three main modules: *Feature Construction*, *Clustering*, and *Cluster Labeling*. The first module builds a feature vector based on WEs for each relation candidate; for doing this, first it must construct a WE dictionary by using a large collection of unannotated texts, in our case extracted from Wikipedia. After that, the K-means clustering is applied to the set of feature vectors obtained by the first module; eventually, a last module associate a set of labels to each cluster. These modules are described in the remainder of this section.

---

<sup>1</sup>Freely available at <https://code.google.com/archive/p/word2vec/>.



**Fig. 1** Architecture for the construction of the WE dictionary

### 3.1 Feature Construction

The steps necessary to build the WE dictionary are depicted in Fig. 1. The first module normalize the input texts and inputs them to the *word2vec* module [2].

#### 3.1.1 Training Set Construction

As mentioned above, we extracted the data set needed to built the WEs dictionary from Wikipedia. More precisely, we considered Italian articles belonging to Medicine, Biology and pharmaceutical categories. In fact, all Wikipedia articles are tagged with at least one category and all categories are organized in a tree-shaped hierarchy and it is possible to identify all the pages related to a specific category and its subcategories. The selection of the pages has been automatically performed using CatScan v3.0,<sup>2</sup> a web tool constructing the list of all Wikipedia articles belonging specific topic. The selection can be refined choosing language, category and depth of the subcategories, that is the number of subcategory of each article to be scanned and included in the results. In our case we have searched all articles in Italian language that belongs to Medicine, Biology and Pharmaceutical category, with a depth equals to five, obtaining a list of 171,877 Wikipedia pages. All the articles in this list can be exported and saved in a standard XML format, using the Wikipedia export tool.<sup>3</sup> We performed this extraction on the 30th of January 2016.

<sup>2</sup><https://tools.wmflabs.org/catscan2/catscan2.php>.

<sup>3</sup><https://en.wikipedia.org/wiki/Special:Export>.

### 3.1.2 Text Normalization

To extract only the plain text from the XML articles exported in the preceding step we used the Wikipedia extractor tool,<sup>4</sup> a python script that generates plain text from Wikipedia XML dumps, discarding any other information or annotation, including images, tables, references and lists.

Further processing is necessary before the plain text can be input to *word2vec*. We therefore implemented a few python and perl scripts, deleting numbers, punctuation, parentheses and all non textual symbols, leaving only the words. After lower-casing, text has been lemmatized by means of TextPro v2.0 suite [7], a Natural Language Processing tool that supports Italian language. Such operation is necessary to reduce data sparsity, also considering that Italian is a morphologically rich language. After deleting stop words,<sup>5</sup> the text has been input to the following module.

### 3.1.3 Word Embedding Dictionary Construction

The construction of the WE dictionary has been performed by means of *word2vec* [2, 8],<sup>6</sup> a tool providing an efficient implementation of the continuous bag-of-words architecture [2] for computing WEs, used in next module of our system. For the feature vectors length we chose 500, which is the default choice, and set the minimum word count to 3, to exclude the less frequent words from the dictionary, obtaining a set of 260, 680 vectors.

### 3.1.4 Input Preprocessing

The system proposed in this paper, without the unsupervised training phase necessary for the construction of WEs, is structured in several modules, depicted in Fig. 2.

The input data set is composed by 989 Electronic Medical Records (EMR) written in Italian and is the same as in [1], where the entity recognition module we apply here is also described. An example of the output it produces for the following portion of text:

*Persiste la nausea con conati di vomito, singhiozzo e malessere generale, digiuno.*<sup>7</sup>

is reported in the following,

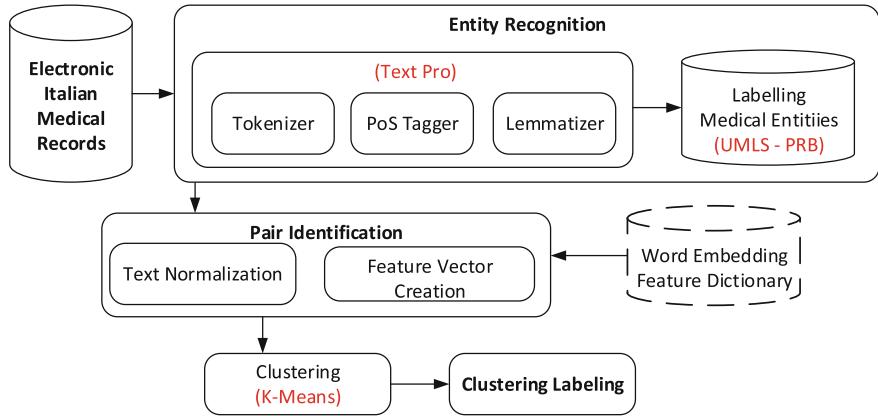
*Persiste la <ent1>nausea</ent1> con <ent2>conati di vomito</ent2>, <ent3>singhiozzo</ent3> e <ent4>malessere generale</ent4>, <ent5>digiuno</ent5>*

<sup>4</sup>[http://medialab.di.unipi.it/wiki/Wikipedia\\_Extractor](http://medialab.di.unipi.it/wiki/Wikipedia_Extractor).

<sup>5</sup>We used the Italian stop word list provided by Lucene and available at [https://lucene.apache.org/core/4\\_4\\_0/analyzers-common/org/apache/lucene/analysis/it/ItalianAnalyzer.html](https://lucene.apache.org/core/4_4_0/analyzers-common/org/apache/lucene/analysis/it/ItalianAnalyzer.html).

<sup>6</sup>The software is freely available at <https://code.google.com/p/word2vec/>.

<sup>7</sup>Nausea persists with vomiting, hiccups, and general malaise, fasting.



**Fig. 2** System architecture for Clustering Labeling. UMLS stay for *Unified Medical Language System*, while PRB for *Pharmaceutical Reference Book*

Token	Id	EoS	Pos tag	Lemma	Entity Type
persiste	146	-	VI	persistere	
la	147	-	SS	det	
nausea	148	-	SS	nausea	B-MED
con	149	-	E	con	
conato	150	-	SP	conato	B-MED
di	151	-	E	di	I-MED
vomito	152	-	SS	vomito	I-MED
,	153	-	XPW	,	
singhiozzo	154	-	SS	singhiozzo	B-MED
e	155	-	C	e	
malessere	156	-	SS	malessere	B-MED
generale	157	-	AS	generale	I-MED
,	158	-	XPW	,	
digiuno	159	-	SS	digiuno	B-MED
.	160	<eos>	XPS	full_stop	

**Fig. 3** Output produced by the entity recognition module for the example sentence reported in the text. The medical entities have been labeled with sub-categories. <eos> stays for “End of Sentence”

while the complete details are reported in Fig. 3, adopting the BIO-Notation, where the prefix B—before the type of entity marks its *beginning* and I—stands for *inside*. The third notation, that is O—for *outside* is never used here, because we consider that entities are always composed by adjacent words.

Then we construct all possible relation candidates by considering all pair of entities occurring in the same sentence: in this way we obtain 77,730 items. Last, we normalize the text corresponding to each entity exactly as described in Sect. 3.1 for the WE construction. In the example above we have five entities, and therefore we obtain the following ten entity pairs:

1. (nausea, conato di vomito)
2. (nausea, malessere generale)
3. (nausea, singhiozzo)
4. (nausea, digiuno)
5. (conato di vomito, malessere generale)
6. (conato di vomito, singhiozzo)
7. (conato di vomito, digiuno)
8. (singhiozzo, digiuno)
9. (singhiozzo, malessere generale)
10. (malessere generale, digiuno)

Note that the order in which entities occur in the text is preserved.

### 3.1.5 Feature Vector Creation

For every entity pair we then construct a Feature Vector (FV) starting from the WE of each word involved. Each entity can be composed by one or more words, as for example *conati di vomito*: in this case, for each entity, we take the average among the WEs of the words composing the entity. As we considered a WE size of 500, we then obtain a FV of the same size for each entity. To build the FV associated to the entity pair we then concatenate the FVs of the two entities, obtaining a size of 1,000.

In this way, each FV can be partitioned in two parts: the former half corresponds to the first entity in the pair, the latter to the other. Such partition is consistently maintained during the whole processing. Also in the computation of centroids in the k-means clustering algorithm, the former half of each centroid derives from the average of the former half of the involved FVs and then corresponds to the first entity. Correspondingly, the latter half of each centroid vector only depends on the second entity of each involved pair.

The choice of the cluster to which a given item is assigned is based on the cosine similarity. Its computation can be divided in three parts: the dot product of the part of the two FVs corresponding to the first entity, the same for the second entity and eventually the normalization with respect to the whole FV. In other words, the evaluation of the cosine similarity is based on a trade-off between how similar are the first and the second entities in each pair. We will discuss later how such behaviour affects the phase of cluster labeling.

## 3.2 Clustering

The clustering algorithm is then applied to the FV data set by means of the *C Clustering library* [9], a fast C implementation of the k-means algorithm. As the k-means is characterized by a random phase due to the initial choice of the seeds, we repeated

each run 10 times, always choosing the best solution. We considered the Euclidean distance, choosing a number of cluster equals to 40, which seemed a reasonable number from the experiments in [1].

### 3.3 Cluster Labeling

To label each cluster we ordered the pairs in each cluster according to its cosine similarity from the cluster centroid: the first four pairs are then chosen to characterize the cluster. As discussed above, these pairs represent a good compromise between considering the closeness of the first and the second entities in each pair. In other words, they represent actual entities pairs which are similar to the (abstract) cluster representative, corresponding to the centroid.

For example in Cluster 11 we have the following four nearest pairs to the centroid: (*retching, pain*)—(*retching, hiccup*)—(*paraesthesia, hypoesthesia*)—(*inappetence, polydipsia*). We note that three pairs have all the entities involved in the relation that are a sign or a symptom of a pathological condition. Almost all the other members of Cluster 11 have in the relation two entities which are signs or symptoms of pathological conditions, like *nausea, pain, pallor, sialorrhea*, etc. This is property of all clusters.

## 4 Discussion of Preliminary Results

In this section we discuss the results obtained by our system. As we do not have a ground truth to quantitatively assess our results, we will discuss them from a qualitative point of view, trying to highlight their strengths and weaknesses. In Tables 1, 2, 3, 4 and 5, we report a few examples of cluster, which are representative of the general pattern we obtained. For each one of them, the first four rows correspond to the label constructed to our system while the caption gives a brief description of their contents.

**Table 1** Cluster 3: diabetes illness and its symptoms

tetraparesi	piede diabetico
arteriopatia	piede diabetico
piedi	piede diabetico
calcoli	retinopatia diabetico
tireotossicosi	iperglycemia
ipoprotrombinemia	iperglycemia
crepitii	iperglycemia
tracheostomia	ipoglicemia
emiplegia sinistra	iperglycemia

**Table 2** Cluster 7: liver and pancreas diseases and related analyses

cirroso	varici esofagee
pancreatite	colelitiasi
pancreatite acuta	noduli
pancreatite	colestasi anitterica
pancreatite	adenopatia
pancreatite	noduli
epatopatia cronico	colelitiasi
epatopatia	colectomia
pancreatite	laparocele
pancreatite cronico	colelitiasi
pancreatite	colecistite
pancreatite	litiasi
cirroso	paracentesi
pancreatite	pseudocisti
pancreatite cronico	epatologia

**Table 3** Cluster 10: chemicals, drugs, proteins and related exams

piastrinopenia	febbrile
colelitiasi	lombalgia
angioma	lombalgia
iperpiressia	dolore a fianco
idronefrosi	dolore a fianco
troponina	dolore retrosternale
urine	dolore colico
allopurinolo	dolore retrosternale
dorsalgia	lombalgia
discectomia	lombalgia
gonartrosi	lombalgia
angioma	dolore lombosacrale
ortopnea	sensazione di costrizione

Note how clustering based on WEs has been effective in grouping together candidate relations which are semantically uniform. On the other hand, this approach can not discriminate whether a particular instance of an entity pair does or does not correspond to a relation, as for this aspect an analysis of the surrounding text is needed. Furthermore, each relation can be stated or negated: also for deciding this aspect an approach which takes into account an analysis of the sentence in which the candidate relation appear is necessary. On the other hand, this approach is very effective in grouping and labeling relations which have been recognized as such by an approach as the one described in [1].

**Table 4** Cluster 11: symptoms

conato di vomito	dolori
conati di vomito	singhiozzo
parestesie	ipostenia
inappetenza	polidipsia
scialorrea	ipertonia
dispnea da sforzo	sensazione di costrizione
pollachiuria	nicturia
stizzoso	stizzoso
catarro	parestesie
vertigini	pallore
poliuria	polidipsia
capogiro	visus
inappetenza	disorientamento
espettorazione	espettorazione
parestesie	ipoestesia
ottundimento	ectasia
dispnea da sforzo	polimialgia reumatico

**Table 5** Cluster 16: chemicals and drugs

clotiapina	aloperidolo
diltiazem	aloperidolo
mirtazapina	trazodone
meningioma	trazodone
lipoma	trazodone
levofloxacina	aloperidolo
ipopotassiemia	furosemide
pregabalin	aloperidolo
ipoproteinemia	furosemide
losartan	furosemide
ticlopidina	aloperidolo
siadh	furosemide
lorazepam	trazodone
creatinchinasi	furosemide
po2	furosemide
levofloxacina	broncodilatatori

## 5 Conclusions and Future Work

An approach for semantic relation clustering has been discussed and applied to a set of Italian patient records. The results have been discussed, by emphasizing pros and contras of the approach. Interestingly, the approach is completely based on machine

learning techniques and do not require annotated data, but only easily available data set such as Wikipedia in addition to the documents to process. It would therefore be interesting to port it to a new language, possibly different from English, which represents the most widely studied among all languages.

## References

1. Alicante, A., Corazza, A., Isgrò, F., Silvestri, S.: Unsupervised entity and relation extraction from clinical records in Italian. *Comput. Biol. Med.* (2016)
2. Mikolov, T., Corrado, G., Chen, K., Dean, J.: Efficient estimation of word representations in vector space. In: Proceedings of the International Conference on Learning Representations (ICLR 2013), pp. 1–12 (2013)
3. Shalev-Shwartz, S., Ben-David, S.: *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, New York (2014)
4. Wei, C., Peng, Y., Leaman, R., Davis, A., Mattingly, C., Li, J., Wiegers, T., Lu, Z.: Overview of the BioCreative V Chemical Disease Relation (CDR) task. In: Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 154–166 (2015)
5. Suárez-Paniagua, V., Segura-Bedmar, I., Martínez, P.: Word embedding clustering for disease named entity recognition. In: Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 299–304 (2015)
6. Hashimoto, K., Stenetorp, P., Miwa, M., Tsuruoka, Y.: Task-oriented learning of word embeddings for semantic relation classification. In: Proceedings of the 19th Conference on Computational Natural Language Learning, CoNLL 2015, Beijing, China, July 30–31, pp. 268–278 (2015)
7. Pianta, E., Girardi, C., Zanoli, R.: The TextPro tool suite. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08), Marrakech, Morocco, pp. 28–30 (2008)
8. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems (NIPS), pp. 3111–3119 (2013)
9. de Hoon, M.J., Imoto, S., Nolan, J., Miyano, S.: Open source clustering software. *Bioinformatics* **20**, 1453–1454 (2004)

**Part IV**

**Advanced ICT for Medical  
and Healthcare**

# A Robust Zero-Watermarking Algorithm for Encrypted Medical Images in the DWT-DFT Encrypted Domain

Jiangtao Dong and Jingbing Li

**Abstract** In order to protect personal information, numerous works has been done in watermarking field. However, there still leaves some problems to be solved: (1) most of the watermarking methods were processed in the plaintext domains, which leave latent risk of exposing host image information, thus it is needed to encrypt the host image and process the watermarking scheme in the encrypted domain; (2) numerous image encryption methods had been searched, while not all of them can meet the robustness requirements when applied in the encrypted domain; (3) for some special fields of watermarking applications, medical images, for example, image integrity is an important criterion that should be strictly taken into account. Thus, that kind of watermarking methods which applies by modifying the pixel values are not suitable in this situation. In order to achieve information hiding in such kind of images, special techniques which do not change image integrity is needed. (4) By utilizing homomorphic encryption scheme, one can process watermark extraction without decrypting the encrypted watermarked image first, while it cost too much time in image encryption and decryption, the computational speed need to be improved. Based on the points mentioned above, we proposed a robust zero-watermarking scheme in the DWT-DFT encrypted domain, which embeds and extracts watermark without modifying the pixel values. Firstly, we encrypted both original medical image and watermark image. Then, we extract the DWT-DFT low frequency coefficients as encrypted medical images' feature vector. In watermark embedding and extraction phases, we adopt zero-watermarking technique to ensure integrity of medical images. Taking “db2” wavelet transform for example, we conduct the experiments on the visual quality and robustness of our watermarking scheme. Experimental results demonstrate that our algorithm achieves not only good watermarking robustness, but also ideal computation speed in the homomorphic encrypted domain.

---

J. Dong · J. Li (✉)

College of Information Science and Technology, Hainan University,  
Haikou, China  
e-mail: jingbingli2008@hotmail.com

J. Dong

e-mail: jiangtao.dong@hotmail.com

**Keywords** Robustness • Zero-watermarking • Feature vector • Homomorphic cryptosystem • DWT-DFT encrypted domain

## 1 Introduction

Medical image is a special kind of image which contains lots of patients' information, whose security should be seriously taken into account. Digital watermarking has been proposed as a possible brick of information protection systems, providing a means to embed a unique code into each copy of the distributed content. However, application of watermarking for multimedia content protection in realistic scenarios poses several security issues [1]. In most previously proposed image watermarking schemes, the embedding and extraction of a watermark often processed in the plaintext domain, which only protect the watermark information while ignoring the security of host image. When watermarking is processed by a third party, for example, an cloud sever, there leaves a latent risk of exposing the raw data.

Cryptosystems play an important role for the afore mentioned problem, i.e., to process the watermark in the encrypted domain. However, not all cryptosystems are appropriate for signal processing in the encrypted domain. Most cryptosystems, such as data encryption standard (DES) and advanced encryption standard (AES), do not retain the algebraic relations among the plaintexts after encryption [2, 3]. A special kind of cryptosystems, the homomorphic cryptosystems, are able to keep the algebraic structure of the plaintext, and thus are particularly suitable for this purpose.

Current solutions based on homomorphic encryption offer provable security at the price of a very high complexity [4]. Here, the bottleneck is the secure embedding module, since all watermarked features have to be encrypted using a costly homomorphic cryptosystem. As an example, in [5] it is reported that on a  $1,024 \times 1,024$  image secure embedding takes about 2 min using a standard personal computer. Hanaa et al. [6] proposed a homomorphic block-based KLT image watermarking which works by segmenting the reflectance component of the host image into blocks by using spiral scan and adding the watermark to every block during the application of the KLT to each block, separately. Li et al. [7] proposed a novel CDWM scheme based on homomorphic encryption and dirty paper pre-coding. They introduced a decryption module before watermark detection to create some nonlinearity and thereby inhibit conventional watermark attacks based on linear operations. Zheng et al. [8] proposed a new signal processing procedure, where the multiplicative inverse method is employed as the last step to limit the data expansion. Bianchi et al. [9] conducted an investigation on the implementation of the discrete Fourier transform (DFT) as well as the fast Fourier transform (FFT) on encrypted signals. A Walsh-Hadamard transform-based image watermarking scheme was proposed in [10], which possesses the character of blind watermark extraction, in both the decrypted domain and the encrypted domain.

In this paper, we propose a robust zero-watermarking algorithm for medical images in the DWT-DFT encrypted domain. In Sect. 2, we introduce the fundamental theory. In Sect. 3, we discuss the zero-watermarking scheme in the DWT-DFT encrypted domain. In Sect. 4, we discuss the robustness of our algorithm under various kinds of attacks based on experimental results, and compared the watermarking methods between plaintext domain and encrypted domain. Finally, we conclude our paper in Sect. 5.

## 2 The Fundamental Theory

In this section, we will give an introduction to the homomorphic cryptosystem and the implementation of DWT and DFT in the encrypted domain. Since DWT cannot resist the geometric attacks, such as scaling and rotation [11], we combine the DWT-DFT hybrid approach to improve the robustness of our watermarking algorithm.

### 2.1 Paillier Homomorphic Encryption

In 1978, Rivest et al. first introduced an idea of homomorphic encryption which permits encrypted data to be manipulated without preliminary decryption [12]. It provides a suitable way for secure signal processing, since it retains the algebraic relations among the plaintext after encryption, so that an algebraic operation on the ciphertexts is corresponding to another operation on the plaintexts. The Paillier cryptosystem [13] is a public key cryptosystem with both the homomorphic property and probabilistic property [14]. It is a partial homomorphic cryptosystem, in which only addition or multiplication homomorphism can be achieved.

The reason why we can use the Paillier cryptosystem to encrypt an image is that the Paillier cryptosystem is a homomorphic cryptosystem. Most of the implementations of secure signal processing are based on the homomorphic properties. And the security of Paillier cryptosystem has been proved. There are also a few kinds of secure signal processing techniques based on the Paillier cryptosystem, such as DWT and DCT in the encrypted domain [15]. Our watermarking scheme is based on such transformations in the encrypted domain.

### 2.2 Non-parametric Probabilistic Models

DWT is a wavelet transform for which the wavelets are sampled at discrete intervals. It provides a simultaneous spatial and frequency domain information of the image. In DWT operation, an image can be analyzed by the combination of analysis

filter bank and decimation operation. The analysis filter bank consists of a pair of low and high pass filters corresponding to each decomposition level. The low pass filter extracts the approximate information of the image whereas the high pass filter extracts the details such as edges. The application of 2D DWT decomposes the input image into four separate sub-bands: low frequency components in horizontal and vertical direction directions (cA), low frequency component in the horizontal and high frequency component in the vertical direction (cV), high frequency component in the horizontal and low frequency component in the vertical direction (cH) and high frequency component in horizontal and vertical directions (cD). cA, cV, cH and cD can also be represented as LL, LH, HL and HH respectively.

The decomposing equation of the Mallat algorithm can be defined as follows:

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j_0, m, n}(x, y) \quad (1)$$

$$W\Psi(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j, m, n}^i(x, y), i = \{H, V, D\} \quad (2)$$

In this paper, we adopt the ‘db2’ wavelet, and process the original medical image within just one layer decomposition in order to maintain sufficient computational speed.

### 2.3 Implementation of DFT in the Encrypted Domain

The Discrete Fourier Transform is a signal analysis theory. The  $M \times N$  medical images’ DFT is done by using:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cdot e^{-j2\pi x u / M} e^{-j2\pi y v / N} \quad (3)$$

$u = 0, 1, \dots, M - 1 ; v = 0, 1, \dots, N - 1$

The  $M \times N$  medical images’ Inverse Discrete Fourier Transform (IDFT) is defined by:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)} \quad (4)$$

$x = 0, 1, \dots, M - 1 ; y = 0, 1, \dots, N - 1$

where  $f(x, y)$  corresponds to the value of the medical image at point  $(x, y)$  and  $F(u, v)$  matches the DFT coefficient at point  $(u, v)$  in frequency domain. Digital images are usually expressed in pixels square, so we set  $M = N$ .

## 2.4 Non-parametric Probabilistic Models

The Logistic map is one of the most famous 1D chaotic maps. It is a simple dynamic nonlinear return with complex chaotic behavior so that we can reproduce it if we have its parameters and initial values. Its mathematical definition can be expressed in the following equation:

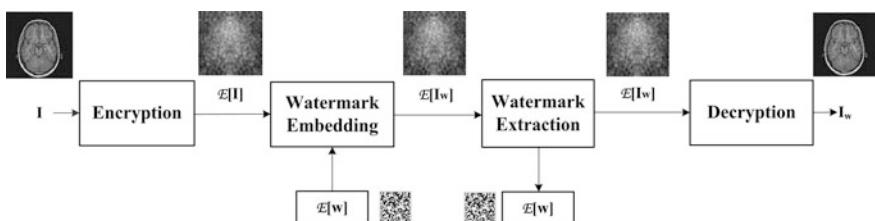
$$x_{k+1} = x_k(1 - x_k) \quad (5)$$

where  $0 \leq \mu \leq 4$  and  $x_k \in (0, 1)$  are the system variable and parameter respectively, and  $k$  is the number of iteration.

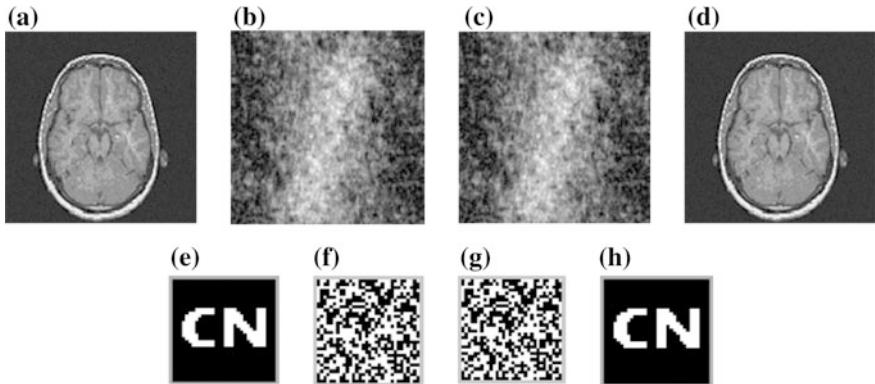
Logistic Map system works under chaotic condition when  $3.569945 \leq \mu \leq 4$ . It can be seen that a small difference in initial conditions would lead to a significant difference of chaotic sequences. These statistical characteristics are the same as white noise, so the above sequence is an ideal secret-key sequence. In this paper, we set  $\mu = 4$ , and the chaotic sequences are generated by different initial values.

## 3 The Algorithm

In this section, we propose a robust image zero-watermarking scheme in the DWT-DFT encrypted domain. As shown in Fig. 1, the original image  $I$  is firstly encrypted by the homomorphic cryptosystem. A scrambled watermark image  $E[w]$  is generated from a Logistic sequence  $Y(j)$ . The watermark embedding is performed in the encrypted domain by a third party, e.g. a cloud server. In order to ensure information security, we have to make sure that the watermark provided to the server is an encrypted binary sequence consists of 0 and 1. Each watermark bit is encrypted separately. Based on the homomorphic property of the cryptosystem, the output of watermark embedding procedure would be an encrypted version of the watermarked image,  $E[I_w]$ . After decryption, the user can acquire the decrypted image which can meets application requirements.



**Fig. 1** Watermarking scheme in the encrypted domain



**Fig. 2** Watermarking in the encrypted domain: **a** the original medical image “mir-1”, **b** encrypted “mir-1” image, **c** watermarked image, **d** decrypted watermark image, **e** original watermark, **f** plaintext domain extraction, **g** encrypted domain extraction and **h** decrypted watermarking robustness performance

### 3.1 Encryption Algorithm of the Original Medical Images

In order to protect the original medical images, we conduct our watermarking algorithm in the encrypted domain. Figure 2 illustrates the encryption scheme for medical images, detailed encryption scheme are described as follows:

- Step 1: Perform DWT on the original medical images to acquire the cA, cH, cV and cD sub-band wavelet coefficients.
- Step 2: Apply DFT on each wavelet sub-band coefficients cA, cH, cV and cD to decompose the sub-band coefficients.
- Step 3: Encrypt all the decomposed sub-band coefficients by dot multiplication operation with a binary matrix C ( $i, j$ ) generated from a Logistic sequence X( $j$ ).
- Step 4: Perform inverse DFT (IDFT) on each coefficients to reconstruct the sub-band.
- Step 5: Perform inverse DWT (IDWT) on the reconstructed coefficients to acquire the encrypted medical images.

### 3.2 Acquire the Feature Vector of Encrypted Medical Images

Taking both the robustness and visual quality into account, we extracted the low-frequency coefficients after applying DWT-DFT hybrid operation. The implementation is given as:

Step 1: Perform DWT on the original medical images to acquire the cA, cH, cV, cD sub-band wavelet coefficients;

$$\{cA, cH, cV, cD\} = \text{DWT2}(I(i,j)) \quad (6)$$

Step 2: Apply DFT to the low frequency coefficients  $cA(i,j)$ ;

$$cA' (i,j) = \text{DFT2}(cA(i,j)) \quad (7)$$

Step 3: Extract the low and middle frequency coefficients of DFT matrix  $cA'(i, j)$ , after binary processing represented in Eq. (8), we can obtain a binary sign sequence of coefficients as the feature vector  $V(j)$ .

$$\text{sgn}(x) = \begin{cases} 1, & x(i,j) \geq 0 \\ 0, & x(i,j) < 0 \end{cases} \quad (8)$$

$$V(j) = \text{sign}(cA' (i,j)) \quad (9)$$

As shown in Table 1. After numerous experiments, we found that the value of DWT-DFT low-frequency coefficients may change after attacking the encrypted image, while the signs of the coefficients still remain unchanged. Let “1” represents a positive or zero coefficient, and “0” represents a negative coefficient. We can obtain the sign sequence of low-frequency coefficients, as shown in the column “Sequence of coefficient signs” in Table 1. After attack, the sign sequence is unchanged, and the Normalized Cross-correlation (NC) is equal to 1.0. Thus we can adopt the coefficient signs as the feature vector of the encrypted medical images.

### 3.3 Watermark Embedding Algorithm in the Encrypted Domain

Medical image is a special kind of image which strictly requires image integrity. Thus, we adopt the zero-watermarking technique to finish watermark embedding. During the process, we embed the watermark by correlating image feature vector and the encrypted watermark  $ew(j)$ , rather than by modifying the image pixel values. Note that, once the key is generated, the watermark is embedded in. The watermark embedding algorithm are described as follows:

- Step 1: Extract the encrypted medical image’s feature vector  $V(j)$  through DWT-DFT transform.
- Step 2: Watermark scrambling. Firstly, we generate a binary logistic sequence  $L(j)$  by using Logistic map; After that, perform hash XOR operation

**Table 1** Change of DWT-DFT coefficients under different attacks for encrypted medical images

Image processing	PSNR (dB)	C(1, 1)	C(1, 2)	C(1, 3)	C(1, 4)	C(1, 5)	Sequence of coefficient signs	NC
Encrypted image without attacks	90.21	7.04 + 0.00i	-2.40 - 0.07i	0.11 + 0.06i	-0.21 + 0.02i	-0.25 - 0.01i	11 00 11 01 00	1.00
Gaussian noise (3 %)	16.02	7.31 + 0. 00i	-2.24 - 0.08i	0.17 + 0.04i	-0.17 + 0.04i	-0.22 - 0.01i	11 00 11 01 00	1.00
JPEG compression (30 %)	19.27	4.98 + 0.00i	-2.09 - 0.07i	0.33 + 0.05i	-0.15 + 0.01i	-0.15 - 0.00i	11 00 11 01 00	1.00
Median filter [3 × 3] (10 times)	24.83	6.96 + 0.00i	-2.44 - 0.07i	0.09 + 0.06i	-0.21 + 0.02i	-0.24 - 0.01i	11 00 11 01 00	1.00
Rotation (clockwise, 2°)	19.78	7.02 + 0.00i	-2.39 - 0.08i	0.07 + 0.07i	-0.18 + 0.01i	-0.27 - 0.00i	11 00 11 01 00	1.00
Scaling (× 0.5)	-	1.76 + 0.00i	-0.62 - 0.02i	0.02 + 0.01i	-0.06 + 0.00i	-0.06 - 0.00i	11 00 11 01 00	1.00
Translation (5 %, down)	15.94	6.98 + 0.00i	-2.39 - 0.07i	0.11 + 0.06i	-0.21 + 0.02i	-0.25 - 0.01i	11 00 11 01 00	1.00
Cropping (5 %, Y direction)	-	6.99 + 0.00i	-2.39 - 0.07i	0.11 + 0.06i	-0.21 + 0.02i	-0.25 - 0.01i	11 00 11 01 00	1.00

between  $L(j)$  and the original watermark  $w(j)$  to acquire the scrambled watermark  $ew(j)$ .

$$ew(j) = w(j) \oplus L(j) \quad (10)$$

Step 3: Generate the  $Key(j)$ . Employ hash XOR operation between  $V(j)$  and  $ew(j)$  to acquire  $Key(j)$ .

$$Key(j) = ew(j) \oplus V(j) \quad (11)$$

### **3.4 Watermark Extraction Algorithm in the Encrypted Domain**

The implementation of watermark extraction in the encrypted domain can be described as:

- Step 1: Perform DWT-DFT on the test image to extract its feature vector  $V'(j)$ .
- Step 2: Employ hash XOR operation between  $V'(j)$  and  $Key(j)$  to extracted the scrambled watermark  $ew'(j)$ .
- Step 3: Watermark recovery. Decrypt the watermark by applying XOR operation between  $ew'(j)$  and the binary Logistic sequence  $Y(j)$ , so we can get the decryption version of extracted watermark  $w'(j)$ .

### **3.5 Watermarking Evaluation**

We use the peak signal-to-noise ratio (PSNR) value to evaluate the visual quality of the watermarked images, and the robustness of the scheme is measured by the Normalized Cross-correlation (NC), which can be described as follows:

$$NC = \frac{\sum_i \sum_j W(i,j)W'(i,j)}{\sum_i \sum_j W^2(i,j)} \quad (12)$$

After detecting  $W'(i,j)$ , compute the NC values between  $W(i,j)$  and  $W'(i,j)$  to determine whether the watermark is embedded in.

The Peak Signal to Noise Ratio (PSNR) is used for measuring the distortion of the watermarked image, which is defined as:

$$PSNR = 10 \lg \left[ \frac{MN \max_{i,j} (I(i,j))^2}{\sum_i \sum_j (I(i,j) - I'(i,j))^2} \right] \quad (13)$$

where  $I(i,j)$  and  $I'(i,j)$  denote the pixel grey values of the coordinates  $(i,j)$  in the original images and the watermarked images respectively; M, N represents the image row and column numbers of pixels respectively.

## 4 Experiments

In our experiment, we select the tenth slice of one medical volume medical data as the original medical image ( $128 \times 128$ ) and choose a significant binary image ( $32 \times 32$ ) as the original watermark image. Figure 2a-d shows the original medical image, encrypted medical image. Figure 2e-h shows the original binary image  $W = \{W(i,j) | W(i,j) = 0, 1; 1 \leq i \leq 32, 1 \leq j \leq 32\}$ . The parameters for encrypting the binary watermark images are:  $x_0 = 0.2$ ,  $\mu = 4$ ; and  $x_0' = 0.135$ ,  $\mu' = 4$  for encrypting the medical images respectively. To verify our algorithm, we run the watermarking scheme on Matlab R2014a platform with a computer contains four Intel(R) Core i5-4590 CPUs at 3.30 GHz.

Experimental results are illustrated in Table 2, in which we compared the robustness performance of watermark attacks in the plaintext and encrypted domains. From the table, we can observe that the watermarking scheme in the encrypted domain achieves good robustness performance which is close to that in the plaintext domain.

To test the computational cost of our algorithm, we conduct our experiment in eight selected images. In the case of embedding a  $32 \times 32$  watermark into a  $128 \times 128 \times 8$  bits' host image, the encryption time is about 0.2 s and the

**Table 2** PSNR and NC values under watermark attacks in the DWT-DFT encrypted domain

Attacks	Parameters	PSNR		NC	
		Plaintext domain	Encrypted domain	Plaintext domain	Encrypted domain
Gaussian noise	5 %	14.68	13.98	0.71	0.88
JPEG	20 %	21.50	19.23	0.96	1.00
Median filter	[ $3 \times 3$ ], 10 times	24.98	24.84	0.90	1.00
Rotation	$4^\circ$	20.59	18.84	0.94	0.88
Scaling	$\times 0.8$	—	—	0.88	0.88
Translation	6 %	14.83	15.69	0.36	0.81
Cropping	10 %	—	—	0.61	0.81

decryption time is about 0.08 s. And the execution time of encrypted domain watermark embedding is about 0.25 s and the extraction time is about 0.05 s in the encrypted domain.

## 5 Conclusions

Most of the existing watermarking schemes were designed to embed the watermark information in the plaintext domain, which leaves a latent risk of exposing the host image. In this paper, we proposed a robust watermarking scheme in the DWT-DFT encrypted domain. The main contributions are listed as follows.

- (1) The proposed watermarking algorithm in the DWT-DFT encrypted domain is robust against watermark attacks in the encrypted domain.
- (2) We adopted the zero-watermarking technique to embed the watermark by correlating image feature vector with a logistic sequence, rather than by changing the image pixel values, thus it can be used in special conditions like medical images etc. that strictly requires image integrity.
- (3) By utilizing our algorithm, the computational speed can easily meets practical requirements.

**Acknowledgments** This work was supported by National Natural Science Foundation of China (No.61263033), International Science and Technology Cooperation Project of China (NO. KJHZ 2015-04) and the Institutions of Higher Learning Scientific Research Special Project of Hainan province, China (NO. Hnkyzx2014-2).

## References

1. Bianchi, T., Piva, A.: Secure watermarking for multimedia content protection: a review of its benefits and open issues. *Sig. Process.* **30**(2), 87–96 (2013)
2. Li, S., Chen, G., Cheung, A., Bhargava, B., Lo, K.-T.: On the design of perceptual MPEG video encryption algorithms. *IEEE Trans. Circ. Syst. Video Technol.* **17**, 214–223 (2007)
3. Zhang, G., Liu, Q.: A novel image encryption method based on total shuffling scheme. *Opt. Commun.* **284**, 2775–2780 (2011)
4. Rial, A., Deng, M., Bianchi, T., Piva, A., Preneel, B.: A provably secure anonymous buye-seller watermarking protocol. *IEEE Trans. Inform. Forensics Sec.* **5**(4), 920–931 (2010)
5. Deng, M., Bianchi, T., Piva, A., Preneel, B.: An efficient buyer-seller watermarking protocol based on composite signal representation. In: Proceedings of 11th ACM Workshop Multimedia and Security, pp.9–18. Princeton, NJ (2009)
6. Abdallah, Hanaa A., Faragallah, Osama S., Elsayed, Hala S., et al.: Robust image watermarking method using homomorphic block-based KLT. *Optik* **127**(4), 2374–2381 (2016)
7. Li, Z., Zhu, X., Lian, Y., et al.: Constructing secure content-dependent watermarking scheme using homomorphic encryption. *IEEE Int. Conf. Multimedia Expo*, 627–630 (2007)

8. Zheng, Peijia, Huang, Jiwu: Discrete wavelet transform and data expansion reduction in homomorphic encrypted domain. *IEEE Trans. Image Process.* **22**(6), 2455–2468 (2013)
9. Bianchi, T., Piva, A., Barni, M.: On the implementation of the discrete Fourier transform in the encrypted domain. *IEEE Trans. Inf. Forensics Secur.* **4**(1), 86–97 (2009)
10. Zheng, P., Huang, J.: Walsh-Hadamard transform in the homomorphic encrypted domain and its application in image watermarking. In: Proceedings of 14th Information Hiding Conference, pp. 240–254 (2012)
11. Kang, X., Huang, J., Shi, Y., Lin, Y.: A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression. *IEEE Trans. Circ. Syst. Video Technol.* **13** (8), 776–786 (2003)
12. Rivest, R.L., Adleman, L., Dertouzos, M.L.: On data banks and privacy homomorphisms. *Found. Secure Comput.* **4**(11), 169–180 (1978)
13. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. *Advances in Cryptology—EUROCRYPT' 99*, pp. 223–238. Springer(1999)
14. Goldwasser, S., Micali, S.: Probabilistic encryption. *J. Comput. Syst. Sci.* **28**(2), 270–299 (1984)
15. Failla, P., Sutcu, Y., Barni, M.: Esketch: a privacy-preserving fuzzy commitment scheme for authentication using encrypted biometrics. In: Proceedings of the 12th ACM Workshop on Multimedia and Security, ACM, pp.241–246 (2010)

# A Touchless Visualization System for Medical Volumes Based on Kinect Gesture Recognition

Ryoma Fujii, Tomoko Tateyama, Titinunt Kitrungrotsakul,  
Satoshi Tanaka and Yen-Wei Chen

**Abstract** The purpose of this study is to construct a system for surgical assistance by touchless interactions. In the clinical site, surgeons usually need to use some contacting devices to display or visualize medical images and check the anatomic information of the patient. However, after operating the visualization device, re-sterilization is necessary in order to maintain hygiene. Though some touchless surgery support systems using Kinect have been developed, their functions are not enough for surgical support. In this paper, we develop a new system, which can visualize 3D medical image by simple touchless single-handed interactions.

**Keywords** Computer aided surgery (CAS) • Medical image visualization • Hand form recognition • Touchless interaction • Kinect

## 1 Introduction

With the development of medical imaging technology, higher resolution 3D medical images nowadays are available for assistance of diagnostics and surgery. Since visualization of these 3D medical images, such as volume rendering and surface rendering, can provide more anatomical information, the surgeon usually needs to use some contacting devices such as mouse, keyboard or touch panel to display the medical images during the surgery operation. However, after operating the visualization device, re-sterilization is necessary in order to maintain hygiene, which is an inefficient and un-effective process for surgery. So development of a touchless

---

R. Fujii (✉) · T. Tateyama · T. Kitrungrotsakul · S. Tanaka · Y.-W. Chen  
Graduate School of Information Science and Engineering,  
Ritsumeikan University, Shiga, Japan  
e-mail: is0135hf@ed.ritsumei.ac.jp

Y.-W. Chen  
College of Computer Science and Technology, Zhejiang University,  
Hangzhou, China

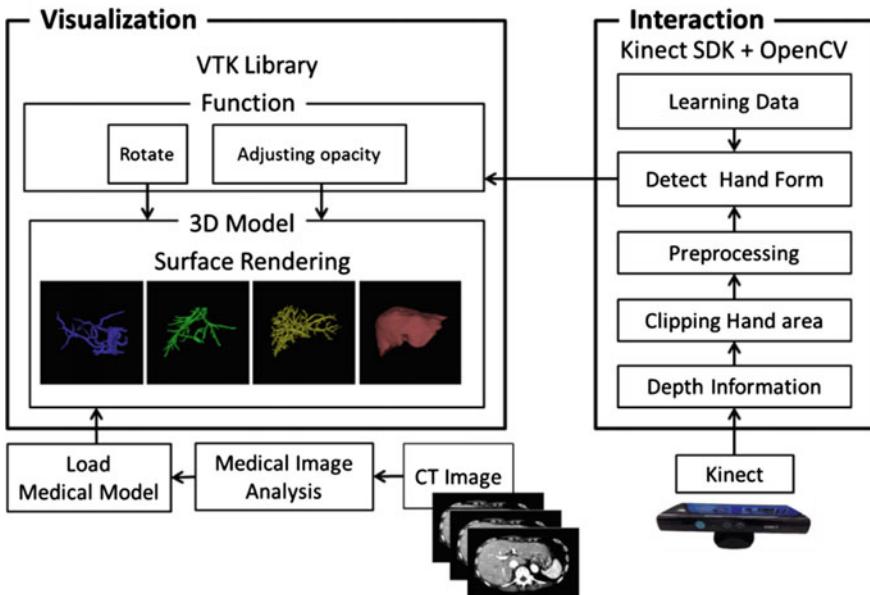
visualization system is an important issue for supporting surgery. In 2010, Microsoft released a low-cost RGB-D camera, called Kinect. Kinect can provide both color image and depth image and it can detect the human actions and human skeleton without any markers. Since it can be used for accurate gesture recognition, Kinect is considered as an ideal solution for touchless interactions. Several touch-less interaction systems have been proposed for visualization of medical images in surgery operation room [1–3]. However, these systems still have some limitations: need two hands for interactions [1] or can only visualize 2D medical image [2]. In this paper, we developed a new touchless visualization system using Kinect for assisting surgery, which can visualize 3D medical images with single-handed interactions.

## 2 Overview of Our Proposed System

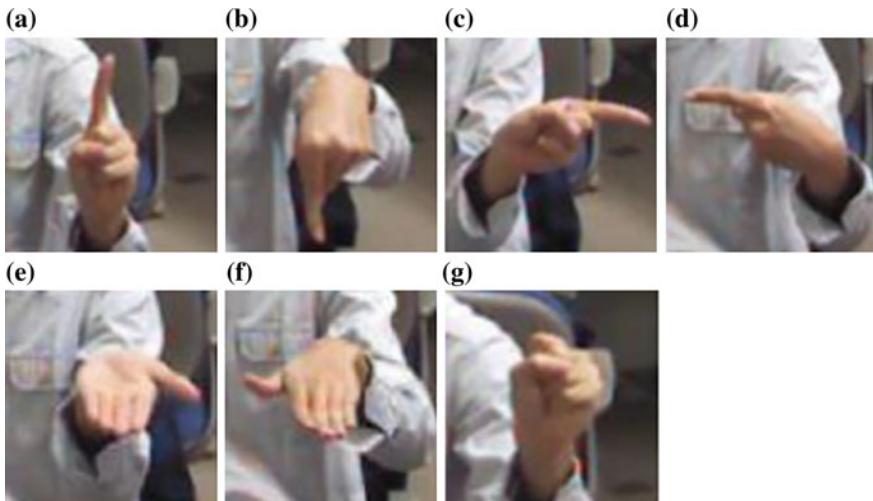
Our proposed system is composed of two modules: interaction module and visualization module. In the interaction module, we designed 7 gestures to control the visualization mode. We use Kinect to obtain a depth image of the hand and recognize the gesture based on a depth image, which is used to control the visualization mode. In the visualization module, 3D surface rendering of the anatomic models are used for fast visualization. There are several visualization modes such as rotation and adjustment of opacity, which are controlled by hand gestures (without touching any device). The diagram of our proposed system is shown in Fig. 1.

## 3 Gesture Recognition

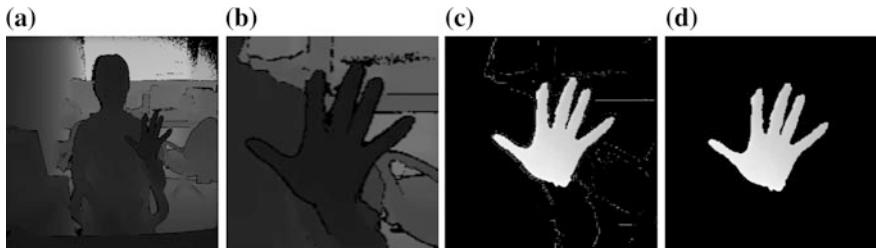
Our system can recognize 7 hand gestures as shown in Fig. 2. They are Finger up, down, right, left, Palm up, down and Grasp. The gesture recognition is based on a depth image obtained by Kinect. Compared with conventional RGB image, the depth image is robust to illumination changes. We first use Kinect to obtain a depth image of the user (Fig. 3a). Then we clip out a square region with the user's right hand in the center and each sides of the square is 36 cm (Fig. 3b). The hand region is defined as  $-30$  to  $5$  cm from surface of the user's right hand as shown in Fig. 4. We extract pixels that have depth value between  $-30$  and  $5$  cm from surface of the user's right hand as a hand image (Fig. 3c). Since the hand image has noise (the depth image obtained by Kinect v2 might contain noise), we use opening operator and median filter to remove the noise and non-hand pixels. In order to normalize the depth value, we subtract a value of  $d$  (the distance between the Kinect and the boundary of hand region as shown in Fig. 4) from the extract hand image (depth image). The location of the hand region is also normalized or aligned by translating the gravity of the hand to the image center (Fig. 3d). We construct a hand image database with 7 gestures generated in this way for gesture recognition. Each gesture has 120 images.



**Fig. 1** Diagram of our proposed system

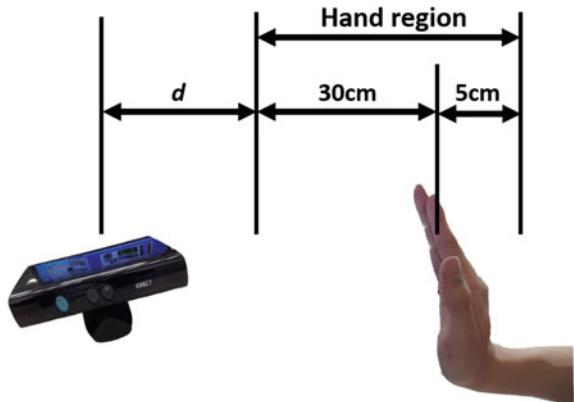


**Fig. 2** 7 hand gestures used in our system. **a** Finger up, **b** Finger down, **c** Finger right, **d** Finger left, **e** Palm up, **f** Palm down, **g** Grasp



**Fig. 3** The depth hand image for gesture recognition. **a** Depth image, **b** Depth hand image, **c** Depth hand image including noise, **d** Depth hand image excluding noise

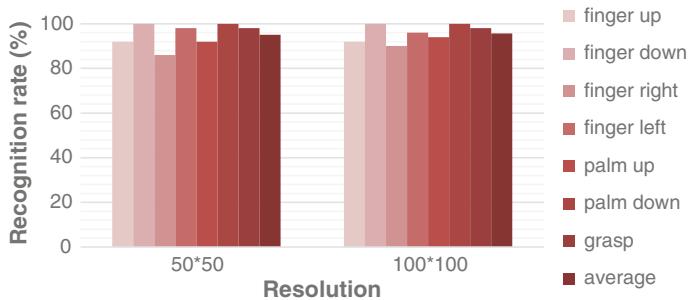
**Fig. 4** Relation of extracted hand region and Kinect position



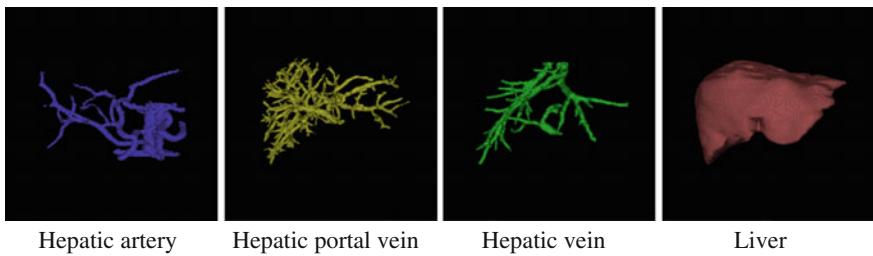
In the hand gesture recognition, we used all pixels' intensity of the hand image (Fig. 3d) as features, and used a  $k$ -nearest neighbor algorithm as a discriminator. The database described above is used for experiments, in which we have 120 images for each gesture. We use 70 images for training and 50 for test. Experiments are performed with both low-resolution ( $50 \times 50$ ) and high-resolution ( $100 \times 100$ ) images. Figure 4 shows recognition rate for each gesture. There are no significant difference between low-resolution ( $50 \times 50$ ) and high-resolution ( $100 \times 100$ ). The recognition rate is significantly affected by the viewing angle. The viewing angle of the input data is different with training data, the recognition rate will be significantly decrease. In our experiments, the average recognition rate is about 95 % (Fig. 5).

## 4 Visualization

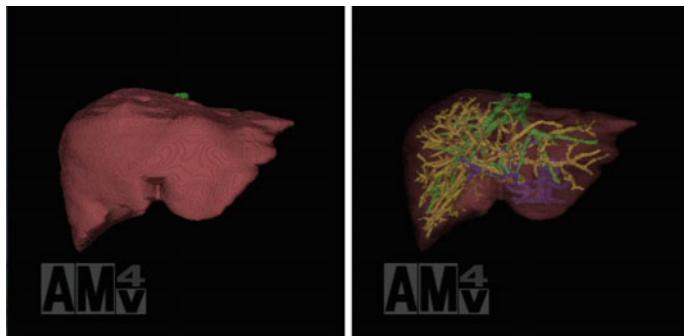
In the visualization, surface models of hepatic artery, hepatic portal vein, hepatic vein and liver parenchyma (Fig. 6) are generated by converting volume data to a triangulated mesh surface by the use of marching cube algorithms [4]. The volume



**Fig. 5** The gesture recognition rate



**Fig. 6** Visualization of liver and its vessels

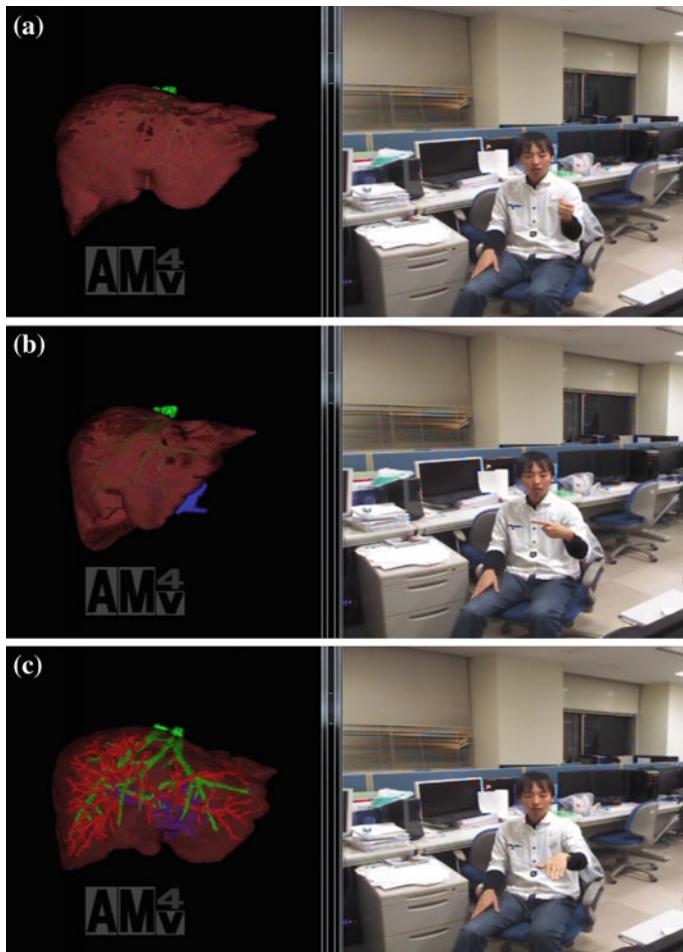


**Fig. 7** Visualization of hepatic region with different opacity. *Left* opaque; *Right* transparent

data are segmented semi-automatically from CT images under the guidance of a physician [5–7]. By visualizing these models together and changing the opacity of the liver, the surgeon can easily recognize the liver geometry, its vessels structures, liver's tumors sizes and locations during the surgery as shown in Fig. 7.

## 5 Touchless Visualization by Single-Handed Interactions

The proposed system adopts 7 gestures. While the user grasps hand or user's right hand is not recognized, the system is in the idle state (Fig. 8a). When the user pointed his index finger into the top, or bottom, or right, or left, the 3D model will be rotated according to the finger direction (Fig. 8b). When the user faced his palm up or down, the opacity of liver will be changed (up: opaque; down: transparent) (Fig. 8c). The surgeon can confirm the information of patient's liver and hepatic vessels based on single-handed interaction (without touching the device) (Table 1).



**Fig. 8** Visualizing operation by single-handed touchless interactions. *Left* patient medical model, *Right* operation user. **a** Operation default mode, **b** Rotation mode, **c** Opacity changing mode

**Table 1** Visualization mode controlled by gestures

Gesture	Visualization mode
Finger up, down, right, left	Rotation of models
Palm up, down	Adjustment of liver's opacity
Grasp	Idling

## 6 Conclusion

We developed a touchless visualization system for medical volume. By using our system, the surgeon can operate and check the information of patient body without touching device. The system can visualize the liver and its vessel structure together and the visualization was controlled by single-handed interactions. We use the depth image obtained by Kinect for gesture recognition and control the visualization mode by hand gestures. In the future, some medical image analysis tools such as segmentation and registration will be incorporated in order to provide more useful information for surgeons.

**Acknowledgements** This work is supported in part by the Grant-in Aid for Scientific Research from the Japanese Ministry for Education, Science, Culture and Sports (MEXT) under the Grant no. 15K16031, no. 15H01130, no. 25280044 and no. 26289069 in part by the MEXT Support Program for the Strategic Research Foundation at Private Universities (2013–2017), and in part by the R-GIRO Research Fund from Ritsumeikan University.

## References

1. Gallo, L. et al.: Controller-free exploration of medical image data: experiencing the Kinect. National Research Council of Italy Institute for High Performance Computing and Networking (2011)
2. Yoshimitsu, K., et al.: Development and initial clinical testing of “OPECT”: an innovative device for fully intangible control of the intraoperative image-displaying monitor by the surgeon. *Neurosurgery* **1**, 46–50 (2014)
3. Ruppert, G.C., et al.: Touchless gesture user interface for interactive image visualization in urological surgery. *World J. Urol.* **30**, 687–691 (2012)
4. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. *Comput. Graph.* **21**, 163–169 (1987)
5. Foruzan A.H., et al.: Segmentation of liver in low contrast images using K-means clustering and geodesic active contour algorithm. *IEICE Trans. Inf. Syst.* **E96-D**, 798–807 (2013)
6. Tomoko, T., et al.: Patient-specific 3D visualization of the liver and vascular structures and interactive surgical planning system. *Med. Imaging Technol.* **31**, 176–187 (2013). in Japanese
7. Chen Y.W. et al.: Computer-aided liver surgical planning system using CT volumes. In: Proceedings of IEEE EMBC2013, Osaka, Japan, pp. 2360–2363 (2013)

# Super-Resolution Technology for X-Ray Images and Its Application for Rheumatoid Arthritis Medical Examinations

**Tomio Goto, Takuma Mori, Hidetoshi Kariya, Masato Shimizu,  
Masaru Sakurai and Koji Funahashi**

**Abstract** Super-resolution techniques have been widely used in fields such as television, aerospace imaging, and medical imaging. In medical imaging, X-rays commonly have low resolution and a significant amount of noise, because radiation levels are minimized to maintain patient safety. In this paper, we propose a novel super-resolution method for X-ray images, and a novel measurement algorithm for treatment of rheumatoid arthritis (RA) using X-ray images generated by our proposed super-resolution method. Moreover, to validate measurement accuracy for our proposed algorithm, we make a model for measurement algorithm about joint space distance using a 3D printer, and X-ray images are obtained to photograph it. Experimental results show that high quality super-resolution images are obtained, and the measurement distances are measured with high accuracy. Therefore, our proposed measurement algorithm is effective for RA medical examinations.

**Keywords** Super-resolution • Rheumatoid arthritis • Joint space distance • Measurement algorithm

---

T. Goto (✉) · T. Mori · H. Kariya · M. Shimizu · M. Sakurai

Department of Computer Science and Engineering,

Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

e-mail: t.goto@nitech.ac.jp

URL: <http://www.splab.nitech.ac.jp/>

T. Mori

e-mail: mori@splab.nitech.ac.jp

H. Kariya

e-mail: kry.hide@splab.nitech.ac.jp

M. Shimizu

e-mail: asimo@splab.nitech.ac.jp

M. Sakurai

e-mail: masaru.sakurai@nitech.ac.jp

K. Funahashi

Department of Orthopaedic Surgery, Nagoya University Hospital,

65 Tsurumai-cho, Showa-ku, Nagoya 466-8560, Japan

e-mail: fkoji@med.nagoya-u.ac.jp

## 1 Introduction

X-ray images are widely used to diagnose a variety of diseases. However, to reduce the patient's exposure to radiation, X-ray dosage is minimized as much as possible. As a result, X-ray images contain a significant amount of noise and resolution is compromised. Thus, it is necessary to increase image resolution and reduce noise. In this paper, we propose a novel super-resolution system for X-ray images that consists of total variation (TV) regularization, a shock filter, and a median filter. In addition, we propose a novel measurement algorithm for the treatment of RA, using X-ray images generated by our proposed super-resolution system.

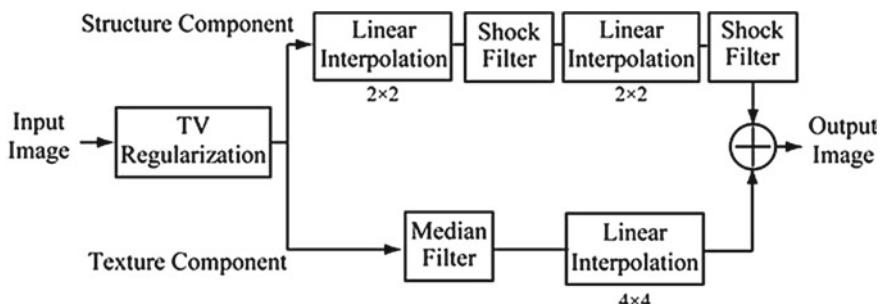
## 2 Super-Resolution System

Super-resolution is a technique for increasing the resolution of an enlarged image by generating new high-frequency components. This technique estimates and generates such components from the characteristics of the original signals. In recent years, various super-resolution techniques have been proposed, and most are classified as either reconstruction-based super-resolution [1] or learning-based super-resolution [2].

Resolution for X-ray images, which is used at medical operation, the resolution of one pixel will be 0.15 mm square, and joint space distance will be about from 1.0 to 1.5 mm. Therefore, when the joint space distance is measured with one pixel error, the error will be from 10 to 15 %, improving measurement accuracy is required.

It is possible to solve this problem by utilizing a super-resolution technique. By magnifying segmented images by a multiple of  $4 \times 4$ , a resolution of a pixel will be 0.0375 mm, thus the error can be suppressed from 2.5 to 3.75 %. Also by utilizing the shock filter, clearer edges in images will be obtained, thus it is easy to select edges.

A block diagram of our proposed super-resolution system is shown in Fig. 1. Each of the nonlinear filters is explained in the following sections.



**Fig. 1** Block diagram of our proposed super-resolution system

## 2.1 Total Variation Regularization

As shown in Fig. 1, the TV regularization decomposition [3–6] is performed as follows [7]. The structure component  $u$  is calculated to minimize the evaluation function  $F(u)$  as shown in Eq. (1):

$$F(u) = \sum_{i,j} |\nabla u_{i,j}| + \lambda \sum_{i,j} |f_{i,j} - u_{i,j}|^2. \quad (1)$$

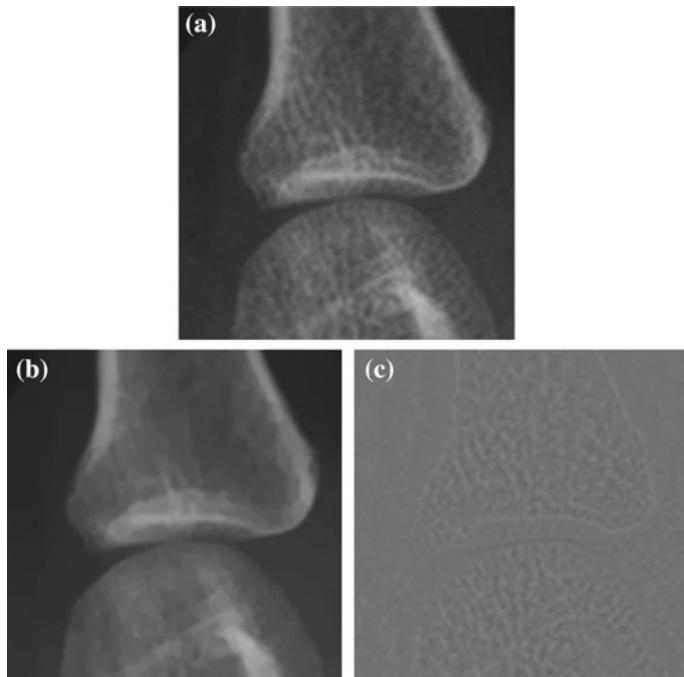
where  $f$  is a pixel value of the input image. The Chambolle's projection algorithm [8] is used to solve the minimization problem.

$$P_{i,j}^{(n+1)} = \frac{P_{i,j}^{(n)} + (\frac{\tau}{\lambda}) \nabla(f + \lambda \operatorname{div} P_{i,j}^{(n)})}{\max\{1, |P_{i,j}^{(n)} + (\frac{\tau}{\lambda}) \nabla(f + \lambda \operatorname{div} P_{i,j}^{(n)})|\}\}}. \quad (2)$$

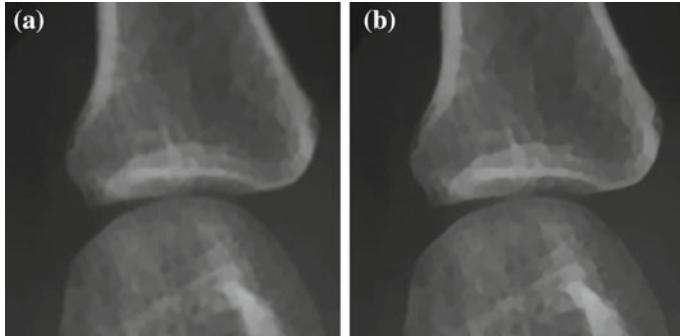
where  $P$  is a pixel value. The texture component  $v$  and the structure component  $u$  are obtained by using the equations in (3).

$$v = \lambda \operatorname{div} P, \quad u = f - v. \quad (3)$$

Figure 2 shows an example of the TV decomposition for an X-ray image.



**Fig. 2** Example of total variation decomposition for X-ray image. **a** Original image. **b** Structure image. **c** Texture image



**Fig. 3** Example of shock filter for X-ray image. **a** Input image. **b** Output image from shock filter

## 2.2 Shock Filter

The shock filter is a nonlinear edge enhancement filter proposed by Osher and Rudin [9] and Alvarez and Mazorra, [10]. The process is achieved by utilizing Eq. (4):

$$u_{i,j}^{(n+1)} = u_{i,j}^{(n)} - \text{sign} \left( \Delta \left( K_\sigma * \Delta u_{i,j}^{(n)} \right) \right) \left| \nabla u_{i,j}^{(n)} \right| dt. \quad (4)$$

where  $u$  is a structure component,  $K$  is a smoothing filter. It is possible to reconstruct steep edges by calculating a simple operation; thus, this filter is suitable for high-speed processing. In addition, several artifacts generated during edge enhancement processing can be controlled successfully, i.e., ringing noise and jaggy noise. Figure 3 shows an input image and output images obtained by utilizing the shock filter.

## 2.3 Median Filter

Most noise is classified as a texture component by utilizing TV regularization. As mentioned previously, X-ray images contain significant noise. Therefore, we propose applying the median filter to the texture components of X-ray images. The median filter sorts nine pixel values in  $3 \times 3$  pixels around the pixel of interest. Next, the filter replaces the fifth pixel value with a new pixel value of interest. This process is applied to all of the texture components.

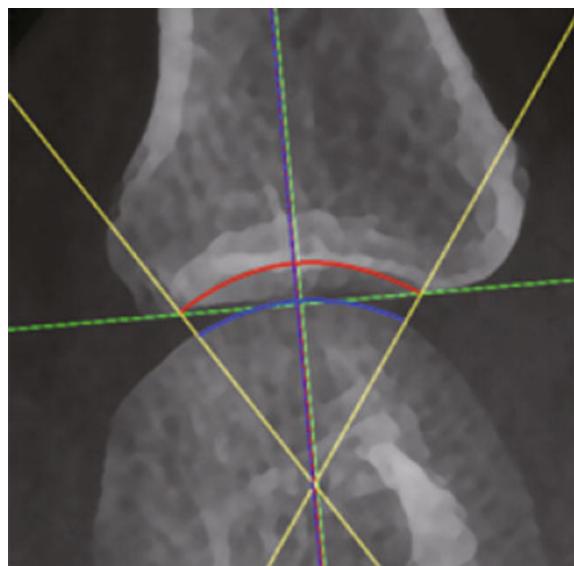
### 3 Measurement Algorithm About Joint Space Distance

Rheumatoid arthritis (RA) is a disease that causes joint inflammation, and most commonly afflicts women between 30 and 50 years of age. As the progress of symptoms, the patient's joint space distance (JSD) will narrow. This change can be observed with X-ray images; however, at present, an accurate measurement method has not been established. Therefore, a more accurate JSD measurement technique is needed. In this paper, we propose two JSD measurement algorithms. Figure 4 shows an example of an output image. In our proposed method, we use an input image, which is magnified by utilizing a super-resolution method. Our algorithm for measurement is shown as follows:

1. Select several points on an edge of an upper bone by clicking mouse button, manually.
2. Set axes corresponding to a joint from selected points automatically.
3. Calculate coordinate values of selected points based on the axes.
4. Calculate a quadratic function by using the least squares method from the calculated fitting function.
5. Calculate a quadratic function by selecting several points on an edge at a lower bone, similarly.
6. Measure the joint space distance from normal lines and integral operation.

We also propose another JSD measurement algorithm, which is calculated by using an area of JSD, and its algorithm for measurement is shown as follows:

**Fig. 4** Application for JSD calculation of X-ray image



1. An integral calculus range in a range of curve p is set, and to calculate the integral calculus  $S_R$  of differences between curve p and curve q.
2. An integral calculus range in a range of curve q is set, and to calculate the integral calculus  $S_B$  of differences between curve p and curve q.
3. The value  $W$  is defined as the distance between 2 points, which is automatically detected when the coordinate axis are set.
4. The JSD value  $D$  is calculated from two values:  $S_R$  and  $S_B$  and averages of the distances of the top and bottom as shown in Eq. (5).

$$D = \frac{S_R + S_B}{2W} \quad (5)$$

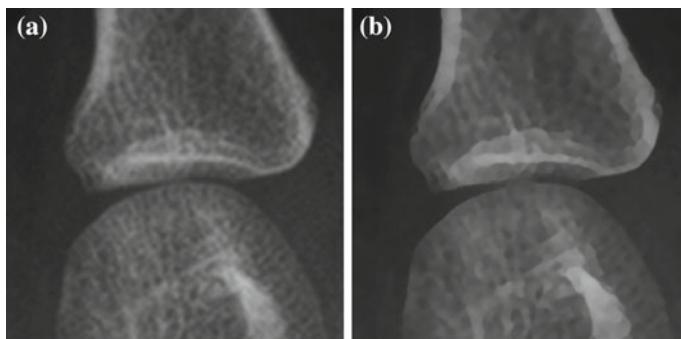
## 4 Experimental Results

### 4.1 Super-Resolution System

Figure 5 shows the experimental results for the super-resolution system, using (a) linear interpolation (Bicubic), and (b) our proposed super-resolution system. It can be seen that the edge of (b) is expanded more clearly than (a).

### 4.2 Measurement System

To validate measurement accuracy in our proposed method, we experiment using 3D modeled images.



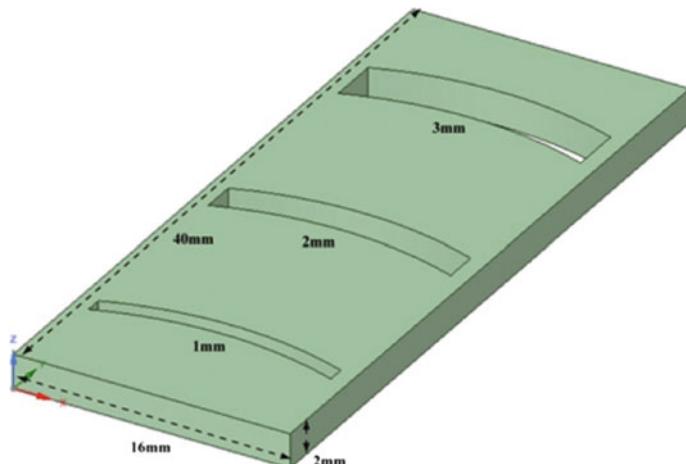
**Fig. 5** Experimental results of super-resolution. **a** Linear interpolation (Bicubic). **b** Proposed super-resolution

#### 4.2.1 3D Modeled Images

We have designed the 3D model in imitation of the joint of the left hand middle finger, which has three joint space distances: 1, 2 and 3 mm as shown in Fig. 6, and have printed out it by using a 3D printer. Then, an X-ray image is obtained to photograph it, we compare the distances between the ideal one and measured one by utilizing our proposed application. The measurement distances are calculated at a mean value and an area by a distance by using Eq. (5).

The experimental condition is shown in Table 1, the designed 3D model, the model made by a 3D printer and its X-ray image are shown in Figs. 7, 8 and 9, respectively. Also, the experimental results for the 3D model is shown in Table 2.

In Table 2, the error between ideal distance and measurement value was 1.3 % for the average distances and 0.4 % for area divided by width, and these error were very small. Therefore, the measurement accuracy was very high quality, and it is possible to verify and monitor the earlier state of RA, objectively.

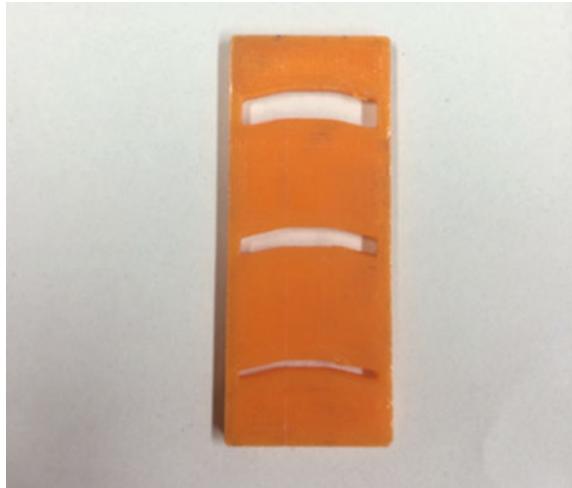


**Fig. 6** Designed 3D model

**Table 1** Experimental parameters

Parameter	Value
Width of joint space distance	3 (mm)
Distance from X-ray to objects	1000 (mm)
Magnified rate	1.5

**Fig. 7** Printed 3D model



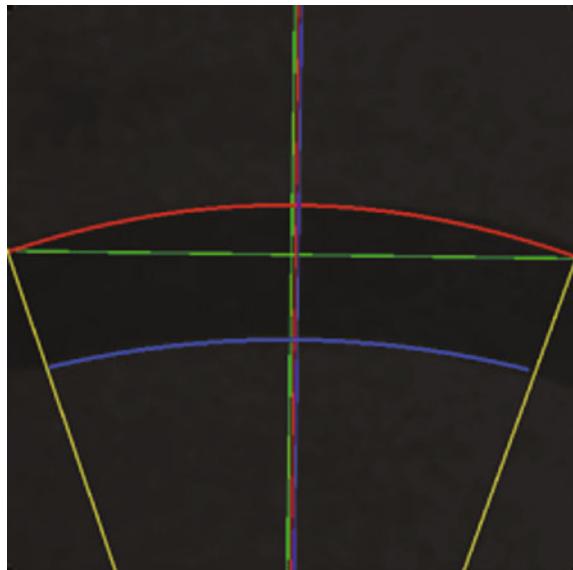
**Fig. 8** X-ray image of printed 3D model



#### 4.2.2 Real X-Ray Images

We will measure the JSD 20 times by using two X-ray images—an input image before the super-resolution, and the super-resolution image. Table 1 shows the variance of measurement results using two X-ray images. The F boundary value is obtained from F distribution table. It is the reference for determining whether there is a significant difference with the variance ratio. In Table 1 the value of the variance ratio is larger than the F boundary value, which implies that there is a significant difference between the average distances before super-resolution and after super-resolution. Therefore, the measurement accuracy can be improved by utilizing our proposed super-resolution system (Table 3).

**Fig. 9** Application for JSD calculation of X-ray image utilizing printed 3D model



**Table 2** Experimental results for X-ray images of printed 3D model

	Average distances (mm)	Area divided by width (mm)
1st	3.0889	3.0047
2nd	3.0725	2.9861
3rd	3.0114	2.9569
4th	3.0006	2.9712
5th	3.0759	3.0171
6th	2.9757	2.9748
7th	3.0254	3.0012
8th	3.2058	2.9890
9th	3.0848	3.0074
10th	3.0252	2.9654
Average	3.0386	2.9874

## 5 Conclusion

In this paper, we have proposed a super-resolution system for X-ray images that utilizes the TV regularization, the shock filter, and the median filter. In addition, we have proposed a novel measurement algorithm for treatment of RA using X-ray images generated by our proposed super-resolution system. Experimental results show that high-quality super-resolution images are obtained from low-quality X-ray images utilizing our proposed super-resolution system. In addition, the results show that

**Table 3** Significant difference test based on variance of before and after super-resolution

	Before super-resolution	After super-resolution
Average (mm)	1.36	1.4733
Variance	0.0517	0.0019
Deviation (mm)	0.2274	0.0436
Number of measurement	20	
Variance ratio	27.2	
F boundary value	2.2	

JSD can be measured more accurately and the progress of RA can be monitored more precisely by using super-resolution images.

Also, we made a model by using a 3D printer to validate our measurement algorithm, and by measuring it, we have validate that measurement accuracy of our proposed method is very high quality.

For further research, we intend to measure JSDs, automatically because several points' selection on edges of upper and lower bones by users is required in our proposed measurement algorithm, and it may include error. Also, we intend to improve the user interface of our proposed measurement system, which doctors use in medical examination and treatment.

## References

1. Singh, A., Ahuja, N.: Single image super-resolution using adaptive domain transformation. In: IEEE International Conference on Image Processing (ICIP), pp. 947–951, Sept 2013
2. Cho, C., Jeon, J., Paik, J.: Example-based super-resolution using selfpatches and approximated constrained least squares filter. In: IEEE International Conference on Image Processing (ICIP), pp. 2140–2144, Oct 2014
3. Rudin, L., Ohser, S., Fetami, E.: Nonlinear total variation based noise removal algorithm. *Phys. D* **60**, 259–268 (1992)
4. Meyer, Y.: Oscillating patterns in image processing and non-linear evolution equation. In: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures, American Mathematical Society, University Lecture Series, vol. 22 (1992)
5. Vese, L.A., Osher, S.J.: Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Comput.* **19** (2003)
6. Aujoil, J.-F., Aubert, G., Blanc-Feraud, L., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vis.* **22**, 71–88 (2005)
7. Goto, K., Nagashima, F., Goto, T., Hirano, S., Sakurai, M.: Super-resolution for high resolution displays. In: IEEE Global Conference on Consumer Electronics (GCCE), pp. 309–310, Oct 2014
8. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**(1), 89–97 (2004)
9. Osher, S.J., Rudin, L.I.: Feature-oriented Image Enhancement using shock filters. *SIAM J. Numer. Anal.* **27**, 910–940 (1990)
10. Alvarez, L., Mazorra, L.: Signal and image restoration using shock filters and anisotropic diffusion. *SIAM J. Numer. Anal.* **31**(2), 590–605 (1994)

# Bayesian Model for Liver Tumor Enhancement

**Yu Konno, Xian-Hua Han, Lanfen Lin, Hongjie Hu, Yitao Liu,  
Wenchao Zhu and Yen-Wei Chen**

**Abstract** Automatic liver lesion enhancement and detection has an essential role for the computer-aided diagnosis of liver tumor in CT volume data. This paper proposes a novel lesion enhancement strategy using Bayesian framework by combining the lesion probabilities based on an adaptive non-parametric model with the processed test volume and the constructed common non-lesion models with prepared liver database. Due to the large variation of different lesion tissues, it is difficult to obtain the common lesion prototypes from liver volumes, and thus this paper investigates a lesion-training-data free strategy by only constructing the healthy liver and vessel prototypes using local patches, which can be extracted from any slice of the test liver volume, and is also easy to prepare the common training non-lesion samples for all volumes. With the healthy liver and vessel prototypes from the test volume, an adaptive non-parametric model is constructed for estimating the lesion possibility, which is considered as the pixel likelihood to lesion region; the common model constructed using the pre-prepared liver database is used to estimate the pixel probability, which is defined as prior knowledge due to the used unvaried model. Finally, the posterior probabilities based on Bayesian theory are achieved for enhancing lesion regions. Experimental results validate that the proposed framework can not only detect almost small lesion regions but also greatly reduce falsely detect regions.

---

Y. Konno (✉) · X.-H. Han · Y.-W. Chen (✉)  
Graduate School of Information Science and Engineering,  
Ritsumeikan University, Shiga, Japan  
e-mail: is0067xk@gmail.com

Y.-W. Chen  
e-mail: chen@is.ritsumei.ac.jp

L. Lin · Y.-W. Chen  
College of Computer Science and Technology, Zhejiang University,  
Hangzhou, China

H. Hu · Y. Liu · W. Zhu  
Radiology Department, Sir Run Run Shaw Hospital, Medical School,  
Zhejiang University, Zhejiang, China

**Keywords** Liver lesion • Prototype • Bayesian framework • Enhancement • Adaptive non-parametric model • CT volume

## 1 Introduction

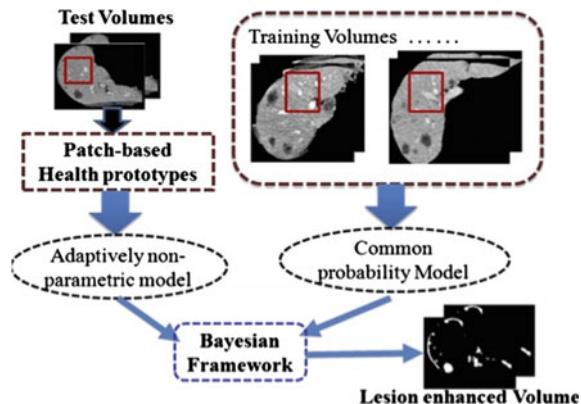
Liver cancer is considered as a major death factor in Japan. In 2012, number of fatality by liver cancer is the 4th most common cause of death by cancer [1]. The survival rate of the early stage 1 of liver cancer is 50.2 %, while those of stage 3 and the last stage 4 are 14.5 % and 6.0 %, respectively. So lesion/tumor detection in early stage is essential for increasing the survival chances of patients. Medical images, such as CT are used as a non-invasive method to detect lesions/tumors (malignant tumors are called cancerous tumors). Recent advancements in medical imaging modalities have enabled the acquisition of high-resolution CT datasets, so that a physician can detect small ones, in addition to large lesions. Owing to the large number of images in medical datasets, it is difficult to manually analyze all images, and useful diagnostic information may be overlooked. Moreover, the diagnoses are mainly based on the physician's subjective evaluation and are dependent on the physician's experience. Therefore, automatic tumor enhancement and detection in CT volumes based on computer vision techniques, which is also known as computer aided detection (CAD), has become one of the major research subjects in the field of medical image analysis.

Until now, many methods have been proposed for lesion enhancement and detection in liver CT images. These methods can be classified as semi-automatic [2, 3] and automatic [4, 5]. Smeets et al. have proposed a semi-automatic level set method, which combines a spiral scanning technique with supervised fuzzy pixel classification [2]. Mala et al. employed wavelet-based texture features in order to train a neural network for use in tumor detection [4]. In the method proposed by Park et al. [5], vessels are removed from liver images and a bimodal histogram is assumed for the intensity distribution of the liver and tumors. The optimal threshold to segment lesions/tumors is determined by a "mixture density" algorithm. Masuda etc. [6, 7] proposed to adaptively enhance the contrast of CT images firstly, and then, applied an expectation maximization and maximization of the posterior marginal (EM/MPM) algorithm by integrating the spatial information for robust tumor segmentation. Foruzan and Chen [8] proposed a sigmoid edge model for tumor segmentation. All the above mentioned methods can locate lesions/tumors that are sufficiently large and have distinct boundaries. Semi-automatic approaches for handling a large number of tumors would need extensive user interactions, and therefore are error prone and tedious.

In this paper, we propose a novel lesion enhancement strategy using Bayesian framework by combining the lesion probabilities based on a non-parametric model with the processed test volume and the constructed common non-lesion models with prepared liver database. The proposed framework mainly includes two models for enhancing lesion regions: (1) the non-parametric model, which is adaptively

generated and dynamical changed according to the selected healthy prototypes of the test volume and can give the likelihood probability for any pixel; (2) the common model, which is constructed from the pre-prepared liver dataset and retains unchanged for any test volume to give prior probability. Final, the posterior probability can be calculated by combining those from the two constructed model with Bayesian framework. The flowchart of the proposed framework is show in Fig. 1. Due to the large variation of different lesion tissues, it is difficult to obtain the common lesion prototypes from liver volumes, and thus this paper investigates a lesion-training-data free strategy by only using the healthy liver and vessel tissues. In order to construct the non-parametric model adaptively from the test volume, a set of liver and vessel tissue prototypes, which are a set of patches of liver and vessel regions, are extracted by randomly displaying any liver slice to user as using local patches. Thus, the non-parametric model is used to calculate the likelihood probability to lesion region. By assuming the liver and vessel pixel intensity to obey a Gaussian distribution, we automatically learn model parameters with Expectation-Maximization (EM) method with the pre-prepared liver and vessel dataset, and thus the generated models with the learned parameters are common for any test volume. By combining the calculated probabilities from the two types of model, not only common information from pre-prepared liver datasets, but also some specific aspect from the test volume can be considered to enhance lesion region. Experimental results show that our proposed framework can correctly detect almost small lesion regions, and at the same time greatly reduce falsely detect regions. The remaining parts of this paper are structured as follows. In Sect. 2, non-parametric probabilistic model adapted to the test volume is described, Sect. 3 introduces the common model constructed using the pre-prepared liver dataset and the proposed Bayesian Model. Section 4 shows experimental results and Sect. 5 concludes the paper.

**Fig. 1** The flowchart of the proposed Bayesian frameworks for liver lesion enhancement

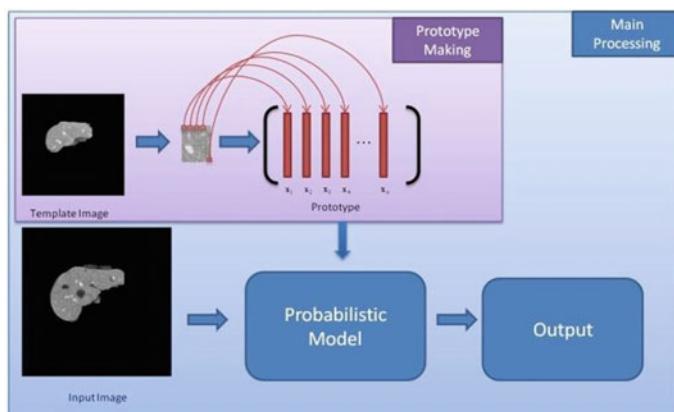


## 2 The Non-parametric Probabilistic Model

In this section, we describe a novel non-parametric probabilistic model for lesion enhancement in order to improve the detection accuracy and reduce the computation cost. The flow of the proposed strategy is shown in Fig. 2. The basic idea is that we extract the health (liver and vessel tissues) prototypes from the test volume and then construct a adaptively non-parametric probabilistic model, which is based on the similarity of input volume patch and the prototypes, to calculate the probability of the lesion pixel (the pixel which does not belong to the pure liver and the vessel) for each voxel. Thus the transformed probability map (the intensity of each voxel is the probability to lesion) can enhance the lesion region.

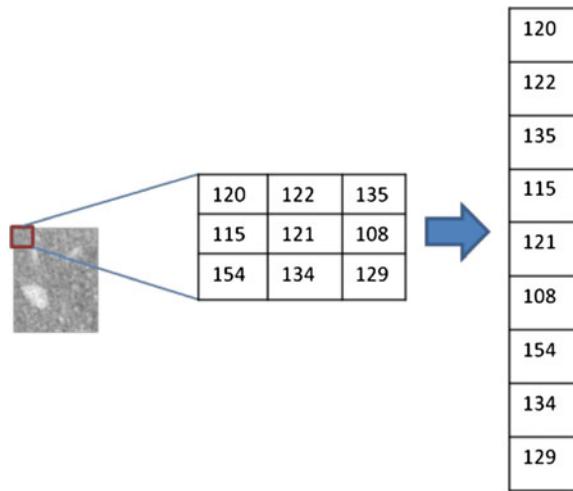
### 2.1 Prototype Preparation

With any segmented test liver volume from the CT data, we randomly selected a slice or several slices, which always include liver and vessel regions, and let user click two points in the selected slice for giving a rectangle region and its two adjacent slices, which only include liver and vessel tissues as a template. Then the liver and vessel prototype set can be generated using the local patches surrounded all voxels in the template image instead of only intensity from the focused voxel. In our experiments, we collect the prototype vector using the intensities of the focused voxel and its 26 neighboring voxels ( $3 \times 3 \times 3$  local patch), and generate totally  $N$  prototypes. Therefore, the dimension of the prototype is  $27 \times 1$ . Figure 3 shows a simple example of prototype generation with  $3 \times 3$  local patch in 2D template case.



**Fig. 2** The flow of the non-parametric probabilistic model

**Fig. 3** The prototype generation procedure from a 2D template



## 2.2 Non-parametric Probabilistic Models

With the generated liver and vessel prototypes, we construct novel non-parametric probabilistic model set for accurately representing the liver and vessel distribution, each of which is from a Gaussian model with a prototype as mean and totally include  $N$  probability model.

Given any test patch  $\mathbf{y}_t$  ( $3 \times 3 \times 3$ ), we can calculate its distances to all prototypes, and choose  $K$  prototypes with the smaller distance. Next, the corresponding models with the  $K$  prototypes as means are used to fit the test patch for calculating probability to liver or vessel ( $P(\mathbf{y}_t|\text{Liver, Vessel})$ ) as shown in the following:

$$P(\mathbf{y}_t|\text{Liver, Vessel}) = \frac{1}{K} \sum_{k=1}^K \left( \exp \left[ -\frac{(\mathbf{y}_t - \mathbf{x}_k)^T \Sigma^{-1} (\mathbf{y}_t - \mathbf{x}_k)}{2} \right] \right) \quad (1)$$

where  $\mathbf{x}_k$  is the  $k$ -th nearest prototype.  $\Sigma$  is the co-variance of the probability models, and for simplicity, we assume all elements of the vector uncorrelated, which thus can simplify the matrix  $\Sigma$  as a diagonal matrix as:

$$\begin{bmatrix} \lambda_{11} & 0 & \dots & 0 \\ 0 & \lambda_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_{nn} \end{bmatrix} \quad (2)$$

where  $\lambda_{nn}$  is the diagonal element of covariance matrix and  $n$  is the dimension of the prototype. Due to non-prior knowledge about the variance of different elements

in prototype vectors, we simply assume the variances of all elements are same, which lead to  $\Sigma$  as a scalar  $\delta^2$  as:

$$\delta^2 = \prod_{i=1}^n \lambda_{ii} \quad (3)$$

Furthermore, since medical volume usually contains noise, the distance of  $\mathbf{y}_i$  to any prototype  $\mathbf{x}_k$  is not zero even the test patch ( $\mathbf{y}_i$ ) is from liver or vessel region. Therefore, taking account of the effect of noise, we rewrite the probabilistic model in Eq. (1) as the following:

$$P(\mathbf{y}_i|\text{Liver, Vessel}) = \frac{1}{K} \sum_{k=1}^K \left( \exp \left[ -\frac{\max[0, ((\mathbf{y}_i - \mathbf{x}_k)^T (\mathbf{y}_i - \mathbf{x}_k) - c)]}{2\delta^2} \right] \right) \quad (4)$$

where  $c = \min(d_{pq})$ .  $d_{pq}$  is the distance between the  $p$ -th and the  $q$ -th prototypes. The probability of the test patch to lesion region (or non-liver and vessel region) can be computed as:

$$P(\mathbf{y}_i|\text{Lesion}) = 1 - \frac{1}{K} \sum_{k=1}^K \left( \exp \left[ -\frac{\max[0, ((\mathbf{y}_i - \mathbf{x}_k)^T (\mathbf{y}_i - \mathbf{x}_k) - c)]}{2\delta^2} \right] \right) \quad (5)$$

### 2.3 *Lesion Enhancement with the Transformed Probability Map*

Given any test volume, we can calculate the probabilities of the test patches centered in all voxels to lesion regions according to Eq. (5). With the calculated probabilities, the input test volume can be transformed to a probability map belonging lesion region, which enhances lesion region and suppress the liver and vessel regions. Thus, it is possible to detect the lesion regions by simply thresholding the probability map.

## 3 Bayesian Model for Liver Lesion Enhancement

### 3.1 *Motivation*

As mentioned in the above section, the non-parametric probabilistic models are constructed with the generated prototypes from the arbitrary slice of the input CT data, and thus the lesion enhancement result will greatly dependent on the selected prototypes. If the un-common region, which has biased distribution from the liver

and vessel, is selected for prototype generation, the lesion region would not be correctly enhanced. Thus, this study proposes to use a common model, which is constructed using the well pre-prepared liver and vessel regions from several CT volumes, to compensate the enhanced result only from the test volume itself; we combine the probabilities of the two types of models using Bayesian theory for final lesion enhancement, called as Bayesian model.

### 3.2 Common Model Construction

With some CT volumes, which include liver organs, the liver and vessel regions can be extracted for preparing the training dataset. Due to the training dataset are selected from several CT samples, and thus it would be more suitable to construct a common model for liver and vessel tissues. Let's assume the common models for liver and vessel obey Gaussian distribution, respectively, the mean  $\mu_i$  and variance  $\Sigma_i$  ( $i = 1, 2$ ) of liver or vessel model can be automatically learned using Expectation-maximization method based on the prepared training data  $\mathbf{X}$ . The likelihood of  $\mathbf{X}$  to fit the learned model can be expressed as:

$$P(\mathbf{X}) = \sum_{i=1}^2 \omega_i N\left(\mathbf{X} | \boldsymbol{\mu}_i, \Sigma_i\right) \quad (6)$$

where  $w_i$  is the weight of  $i$ -th component and  $w_1 + w_2 = 1$ . Via maximizing the above function, we can automatically learn the model parameters  $\boldsymbol{\mu}_i$ ,  $\Sigma_i$ .  $N(\mathbf{X} | \boldsymbol{\mu}_i, \Sigma_i)$  denotes a Gaussian distribution and can be expressed:

$$N\left(\mathbf{X} | \boldsymbol{\mu}_i, \Sigma_i\right) = \frac{\exp[-(\mathbf{X} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{X} - \boldsymbol{\mu}_i)]}{2\pi |\sqrt{\Sigma_i}|}$$

With the learned Gaussian model (common to any liver and vessel region), we can fit any small (liver, vessel or lesion) to them, and obtain the larger value for representing probability of this region belonging to liver or vessel. Given any test patch  $\mathbf{y}_t$ , the fitting probability to the constructed model is calculated as:

$$P_i(i|\mathbf{y}_t) = \frac{\exp[-(\mathbf{y}_t - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{y}_t - \boldsymbol{\mu}_k)]}{2\pi |\sqrt{\Sigma_k}|} \quad (8)$$

$$P(\text{lesion}) = 1 - \max[P_i(i|\mathbf{y}_t)] \quad (9)$$

## 4 Experiments

In order to validate the efficiency of our proposed framework, we applied our proposed methods to abnormal CT volumes. The used CT dataset was released as the MICCAI Liver Tumor Segmentation Challenge 2008 [9], which has size of  $512 \times 512 \times 512$  and the spacing of  $0.72 \text{ mm} \times 0.72 \text{ mm} \times 1 \text{ mm}$ . There are 50 tumors labeled by experts. As a preprocessing, we use an anisotropic diffusion filter to remove the noise contained in the CT volume, which can effectively reduce the noise without significant blurring. The liver is segmented manually from the CT volume. Recently, we have developed several semi-automatic [10, 11] or automatic [12] liver segmentation methods, which can be used as preprocessing in the future.

In order to make a comparison, we also use the conventional GMM (Gaussian Mixture Model) based method, which learns parameters using all voxels in the test volume and takes high computational time, for lesion detection. Figure 4 show the lesion detection results with the conventional GMM method, only the non-parametric probabilistic model and the Bayesian framework. It can be seen that some lesion regions cannot be detected by the conventional GMM method. The final statistical detection results are summarized in Table 1, which manifest that our proposed two methods both can detect all lesion regions with less false positive compared with conventional GMM method.



**Fig. 4** Input image (*left*), GMM result (*center*) and propose method result (*right*)

**Table 1** The final statistical lesion detection result

	False positive	True positive	Total tumor
Conventional method (GMM)	1687	39	50
Proposed method (non-parametric probabilistic model)	364	50	
Proposed method (Bayesian model)	250	50	

## 5 Conclusions

This paper proposes a novel lesion enhancement strategy using Bayesian framework by combining the lesion probabilities based on a non-parametric model with the processed test volume and the constructed common non-lesion models with prepared liver database. The proposed framework not only consider the specific information from the test volume but also take the distribution of training liver and vessel regions as a common aspect for lesion enhancement. Experimental results showed that the proposed framework can both detect almost small lesions and greatly reduce falsely detect regions.

**Acknowledgments** This research was supported in part by the Grant-in Aid for Scientific Research from the Japanese Ministry for Education, Science, Culture and Sports (MEXT) under the Grant No. 15H01130 and No. 15K00253, in part by the MEXT Support Program for the Strategic Research Foundation at Private Universities (2013–2017), and in part by the Recruitment Program of Global Experts HAI OU Program from Zhejiang, China.

## References

1. National Cancer Center, Japan: Center for Cancer Control and Information Services. <http://ganjoho.jp/public/statistics/pub/statistics01.html>
2. Smeets, D., et al.: Semi-automatic level set segmentation of liver tumors combining a spiral scanning technique with supervised fuzzy pixel classification. *Med. Image Anal.* **14**, 13–20 (2010)
3. Hame, Y., et al.: Image analysis for liver tumor ablation treatment planning, hands-on image processing 2009. Robotiker-Tecnalia
4. Mala, K., et al.: Neural network based texture analysis of liver tumor from computed tomography images. *Int. J. Biomed. Sci.* **2**, 33–40 (2006)
5. Park, S.-J. et al.: Automatic Hepatic Tumor Segmentation Using Statistical Optimal Threshold, Computational Science-ICCS2005, vol. 3514, pp. 934–940. Springer, Berlin (2005)
6. Masuda, Y., et al.: Automatic liver tumor detection using EM/MPM algorithm and shape information. *IEICE Technical Report*, vol. 110, pp. 25–30 (2010)
7. Masuda, Y., et al.: Liver tumor detection In CT images by adaptive contrast enhancement and the EM/MPM algorithm. In: Proceedings of IEEE International Conference on Image Processing (ICIP2013), pp. 1453–1456 (2011)
8. Foruzan, A.H., Chen, Y.-W.: Improved segmentation of low-contrast lesions using sigmoid edge model. *Int. J. CARS* (2015). doi:[10.1007/s11548-015-1323-x](https://doi.org/10.1007/s11548-015-1323-x)
9. Deng, X., Du, G.: Editorial: 3D segmentation in the clinic: a grand challenge II—liver tumor segmentation. <http://grand-challenge2008.bigr.nl/proceedings/liver/articles.html>
10. Foruzan, A.H., et al.: Segmentation of liver in low-contrast images using k-means clustering and geodesic active contour algorithms. *IEICE Trans.* **E96-D**, 798–807 (2013)
11. Dong, C., et al.: Simultaneous segmentation of multiple organs using random walks. *J. Inf. Process. Soc. Jpn.* **24**, 320–329 (2016)
12. Dong, C., et al.: Segmentation of liver and spleen based on computational anatomy models. *Comput. Biol. Med.* **67**, 146–160 (2015)

# Fused Visualization with Non-uniform Transparent Surface for Volumetric Data Using Stochastic Point-Based Rendering

**Kyoko Hasegawa, Yuta Fujimoto, Rui Xu, Tomoko Tateyama,  
Yen-Wei Chen and Satoshi Tanaka**

**Abstract** In medical, scientific, and other fields, transparent surface visualization is useful for investigating inner three-dimensional (3D) structures. Such visualization typically involves the use of polygon graphics in which the polygons must be sorted along the line of sight. However, sorting involves considerable computation time for large-scale data. Furthermore, the order of polygons in the sorting can often become indefinite, especially for intersecting surfaces. Therefore, particle-based volume rendering that does not require sorting is proposed as a transparent-rendering method. The proposed method obtains slice images with non-uniform opacity using color and opacity maps similar to volume rendering. The method initially generates the particles, a process it performs only once. It additionally employs particle shuffling, which requires considerably less computation time than particle sorting. To demonstrate the efficacy of the proposed method, we rendered 3D-fused images, including slice–slice and volume–slice images, for medical volumetric data. The

---

K. Hasegawa (✉) · T. Tateyama · Y.-W. Chen · S. Tanaka  
College of Information Science and Engineering,  
Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga, Japan  
e-mail: hasegawa@media.ritsumei.ac.jp

T. Tateyama  
e-mail: tomoko@media.ritsumei.ac.jp

Y.-W. Chen  
e-mail: chen@media.ritsumei.ac.jp

S. Tanaka  
e-mail: stanaka@media.ritsumei.ac.jp

Y. Fujimoto  
Graduate School of Science and Engineering,  
Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga, Japan  
e-mail: is036080@ed.ritsumei.ac.jp

R. Xu  
School of Software, Dalian University of Technology,  
No. 321 Tuqiang Road, Economy and Technology Development Area,  
Dalian, People's Republic of China  
e-mail: xurui@dlut.edu.cn

results show that the performance of the proposed method is satisfactory in cases in which the area of the particle is greater than that of the cell.

**Keywords** Particle-based surface rendering • High-resolution volumetric data • Fused visualization

## 1 Introduction

In medical, scientific, and other fields, transparent surface visualization is used to investigate interior three-dimensional (3D) structures [1, 2]. Typically, this type of visualization uses polygon graphics [3, 4], and the polygons are sorted along the line of sight. However, this sorting involves considerable computation time for large-scale data. In addition, the order of the polygons in the sorting can often become difficult to interpret especially for intersecting surfaces.

A method that requires no sorting and is applicable to large-scale data was recently proposed by Koyamada et al., whose particle-based volume rendering (PBVR) uses tiny particles as rendering primitives [5]. Their method enables natural volume fusion [6]. In this study, we extended PBVR to make it applicable to surfaces and volumes [7]. We refer to this method as stochastic point-based rendering (SPBR).

An important advantage of SPBR is that the 3D-fused visualization of different volume/surface/slice objects becomes possible simply through the merging of particles prepared for each element to be fused. A surface is visualized at a constant opacity by applying a uniform sampling method.

In this research, a non-uniform transparency slice image is rendered with an opacity map using a transfer function of the color and opacity same as volume rendering. We propose a sampling technique to render the same volume. In addition, we render the 3D-fused visualization, including volume-slice images, based on SPBR.

## 2 Stochastic Point-Based Rendering

SPBR renders 3D scalar fields as particle clouds and incorporates both the emission and absorption effects. Particle density depends on the transfer function and is used to evaluate the number of particles to be generated in the volume data. Because the particles can be considered opaque, no processing for visibility sorting is required. SPBR has three processes: particle generation, particle projection onto the image, and ensemble averaging of particle luminosities. The first process stochastically generates particles according to the form of their transfer function. We repeat the particle generation until  $L_R$  statistically independent particle sets are prepared. In the following, we refer to  $L_R$  as the “repeat level.” Such particle projection with the

occlusion effect is executed for each particle set prepared in the first process. As a result,  $L_R$  similar images are created. The second process projects particles onto an image plane. The third process calculates the ensemble average of the  $L_R$  images created in the second process, which realizes transparent images [5].

To visualize a surface, we stochastically generate particles on the surface [8]. The number of particles  $n$  is defined by the following formula:

$$n = \frac{\ln(1 - \alpha)}{\ln(1 - s_p/s_A)} L_R, \quad (1)$$

where  $\alpha$ ,  $s_p$ , and  $s_A$  are the opacity, the area of the particle, and the area of the surface, respectively [7]. The slice plane is rendered in a manner similar to the one described in the above technique.

## 2.1 Transparency Control for Surface

In this subsection, we describe the method for creating particles in a polygonized surface defined with non-uniform opacity. SPBR may be applicable to defined drawing surfaces. For the case of a polygonized surface defined with uniform opacity, the total number of particles can be calculated by replacing the counting sphere with the total area size.

First, we explain the method for achieving uniform sampling of a polygonized surface; i.e., a polygon mesh with uniform opacity. For a polygonized surface, it is easy to calculate the total areas of the constituent polygons, which we use as area size  $s_A$  in Eq. (1). Correspondingly, we regard the total number of particles generated on all the constituent polygons as  $n_{\text{all}}$ .

The sampling is performed such that the generated particles form a two-dimensional (2D) square grid with inter-particle distance  $d$  in each polygon. Then, the particle density becomes  $1/d^2$ , i.e., one particle per square. On the other hand, the particle density should be  $n_{\text{all}}/s_A$ . Therefore, the equation  $1/d^2 = n_{\text{all}}/s_A$  holds, and  $d$  is determined as:

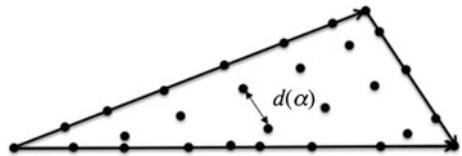
$$d = \sqrt{\frac{(s_A - s_p)}{\ln(1 - \alpha)L_R}}. \quad (2)$$

The sampling involves three steps:

- Step 1 Sampling polygon edges
- Step 2 Sampling inside of polygons
- Step 3 Particle shuffling

By executing step 1 first, duplicative sampling of edges shared by neighboring polygons is avoided. In that step, we simply place particles with distance  $d$  on each

**Fig. 1** Creating particles on the polygon for a non-uniform transparent surface

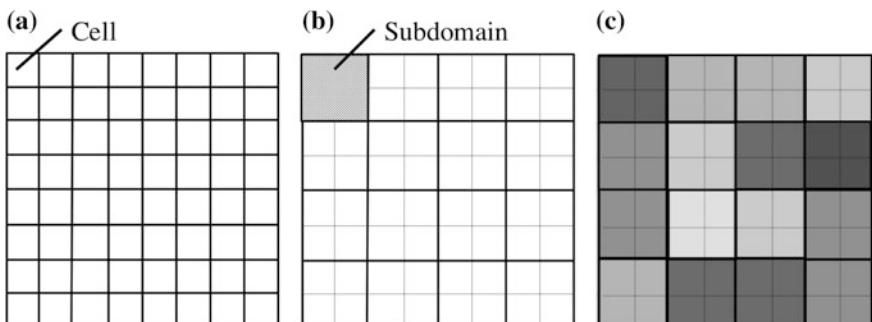


edge of the entire polygon mesh. In step 2, we place particles with distance  $d$  on scan lines for each polygon. The scan lines are placed parallel to an edge at intervals of  $d$ . In step 3, the generated particles that are stored in an array are shuffled. That step ensures the randomness, which is required to apply the probabilistic theory of opacity developed in SPBR. It should be noted that the computation time required for particle shuffling is much shorter than that for particle sorting. The former is proportional to  $n_{\text{all}}$ , while the latter is proportional to  $n_{\text{all}} \log n_{\text{all}}$ . After completing the above three steps, we obtain a particle set that is uniformly distributed and has a desirable statistical property.

In the case of non-uniform transparent surfaces, distance  $d$  is defined with the function of the opacity,  $\alpha$  (Fig. 1). Here, the area of the surface  $s_A$  is defined as the region of a polygon. The opacity on the surface is obtained by projecting onto the volume texture [9], that is the color and opacity map defined as volumetric data, and the distance is defined with Eq. (2) [10].

## 2.2 Slice Transparency Control

In order to visualize the slice plane, the particles are stochastically and uniformly generated on the slice plane. The number of particles is defined as in Eq. (1). A non-uniform transparent slice can be obtained by calculating the number of particles at each cell. Here, the “cell” refers to the cross-section surface of the cube of the volumetric data (see Fig. 2a). The area of surface  $s_A$ —i.e., the area of the cell—is much greater than the area of the particle,  $s_p$ . By uniformly generating the particle at



**Fig. 2** Definition of the cell and subdomain for non-uniform transparent slice image. **a** Cell. **b** Subdomain. **c** Final image

each cell, a slice plane with non-uniform transparency can be created throughout the entire plane.

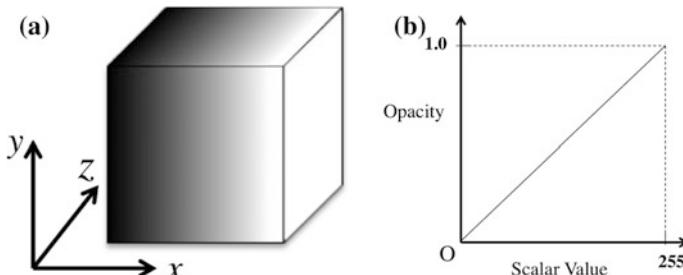
In the case in which the area of particle  $s_p$  is greater than the area of the cell—e.g., high-resolution volumetric data—the number of particles at each cell cannot be accurately calculated. Therefore, the particles cannot be generated in the cell. However, it should be noted that the subdomain consists of several cells, as shown in Fig. 2b. The opacity of the subdomain is defined as the average of the cells. Using the proposed method for a subdomain, the particles are uniformly generated at each cell, and a slice plane with non-uniform transparency can be created throughout the entire plane (see Fig. 2c). The above technique can ensure integrated control of the volume, surface, and slice rendering by using the color and opacity map for SPBR.

### 3 Experiments

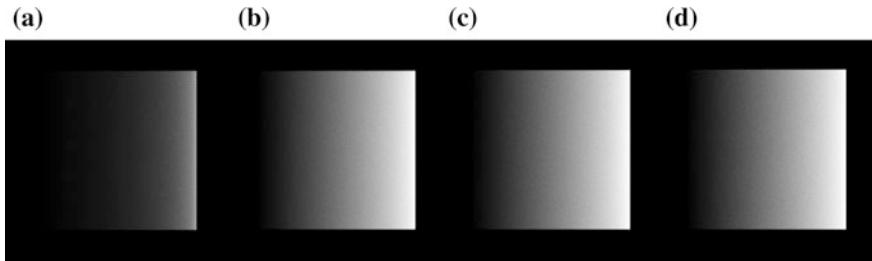
#### 3.1 Test Case for Non-uniform Transparent Slice

Consider the following test-volume data for simplicity. The scalar value  $\rho$  range of  $\rho$  is  $0 \leq \rho \leq 255$ , which linearly increases with  $x$ . The opacity map is defined as in Fig. 3. Here, the color of all particles is white. Figure 4 shows the slice image at  $z = 10$ , and  $L_R = 500$  for the case of different subdomain size. Here, the area of particle  $s_p$  and the area of cell  $s_A$  are 0.996 and 1.0, respectively.

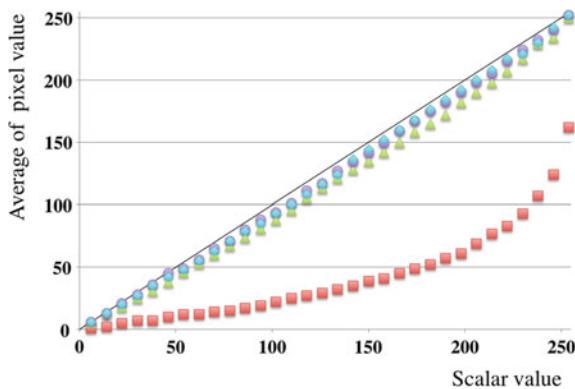
The case of the  $1 \times 1$  cell in the subdomain (see Fig. 4a) shows that this image is darker than the other three slice images owing to an incorrect calculation of the number of particles. The other three cases (Fig. 4b–c) look alike. The results for the number of cells of  $1 \times 1$ ,  $2 \times 2$ ,  $3 \times 3$  and  $4 \times 4$  are plotted in Fig. 5. This figure shows that the cases of  $3 \times 3$  and  $4 \times 4$  behave similarly to the theoretical value. In addition, the least value of the maximum error and average error is for the case of  $3 \times 3$  cells as shown by Table 1. In the case of  $4 \times 4$  cells, the discretization error is deemed to be caused by the expansion of the subdomain.



**Fig. 3** Test-volume data and opacity map. **a** Test-volume data. **b** Opacity map



**Fig. 4** Slice image of the test volume data (xy plane). **a**  $1 \times 1$ . **b**  $2 \times 2$ . **c**  $3 \times 3$ . **d**  $4 \times 4$



**Fig. 5** Pixel value on the slice plane. *Black line* theoretical value, *Red marker*  $1 \times 1$  cell, *Green marker*  $2 \times 2$  cells, *Purple marker*  $3 \times 3$  cells, *Blue marker*  $4 \times 4$  cells

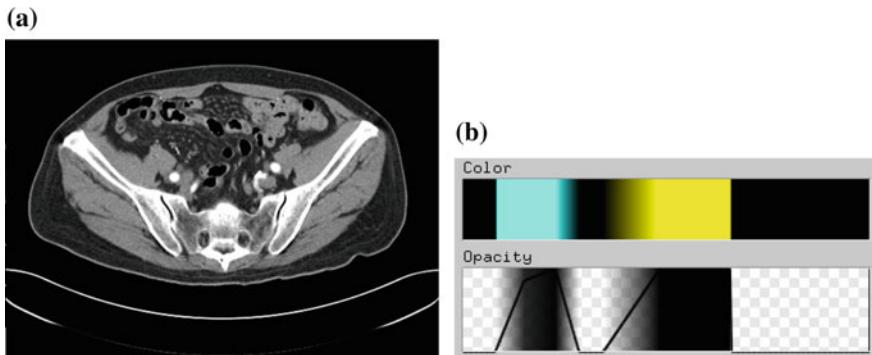
**Table 1** The maximum error and average error for case of each subdomain

Number of cells in the subdomain	$1 \times 1$	$2 \times 2$	$3 \times 3$	$4 \times 4$
Maximum error (%)	85.7	18.4	9.7	11.3
Average error (%)	72.5	10.8	5.5	5.9

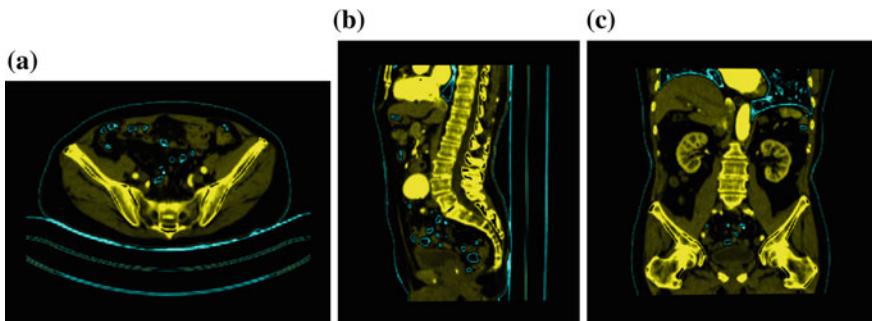
In summary, the results of this study suggest that, for a case in which the area of the particle is greater than that of the cell, a subdomain, with an area that is approximately 10 times the area of particle  $s_p$  will suffice.

### 3.2 Medical Data Experiments

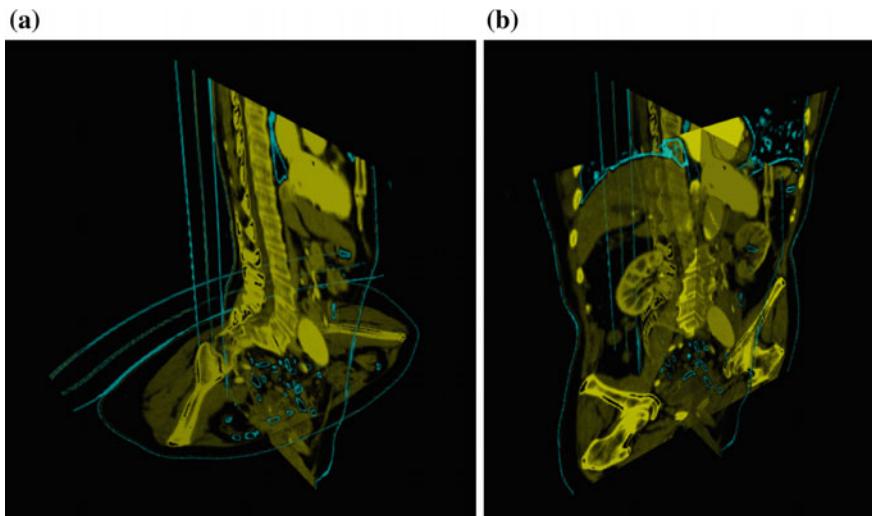
In this section, we present fused visualizations, including slice-slice, surface-slice, and volume-slice visualizations, based on medical data. Throughout this section, the number of cells for a subdomain are fixed as  $3 \times 3$  and  $L_R = 500$ .



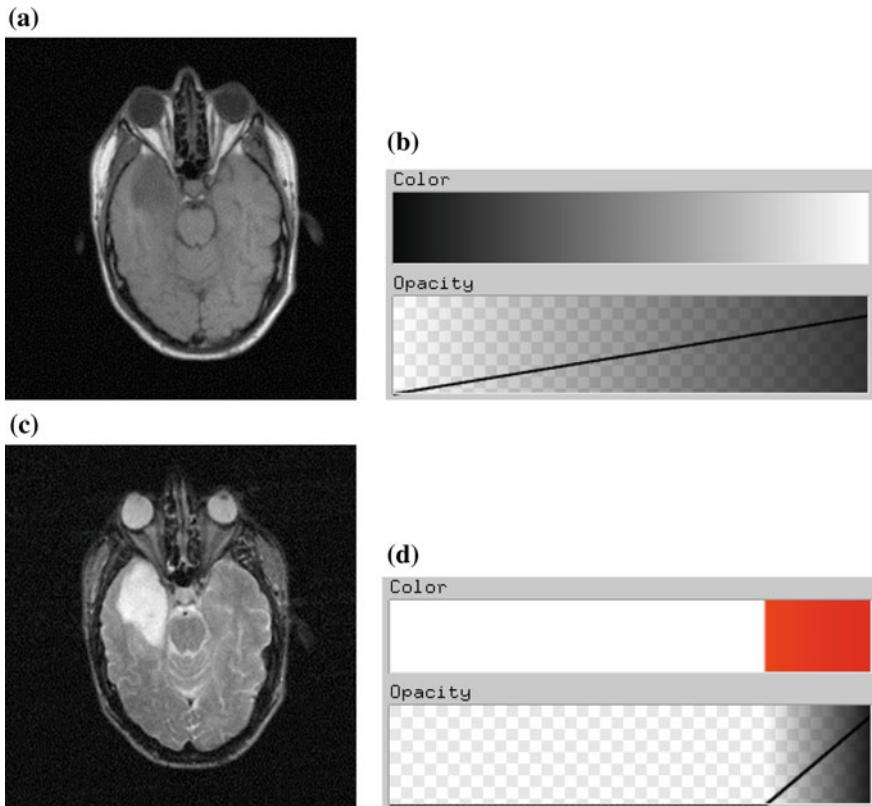
**Fig. 6** Abdominal CT image and color and opacity map. **a** Original CT image. **b** Color and opacity map



**Fig. 7** Result of the slice image visualization. **a** Axial plane. **b** Sagittal plane. **c** Coronal plane



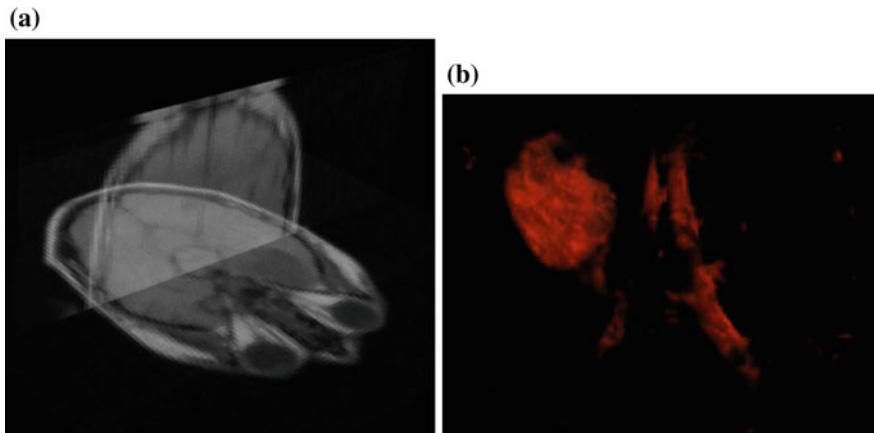
**Fig. 8** Fused visualization of slice rendering. **a** Fused sagittal and axial plane. **b** Fused sagittal and coronal plane



**Fig. 9** Set of the MRI image and color and opacity map. **a** T1-weighted image. **b** Color and opacity map. **c** T2-weighted image. **d** Color and opacity map

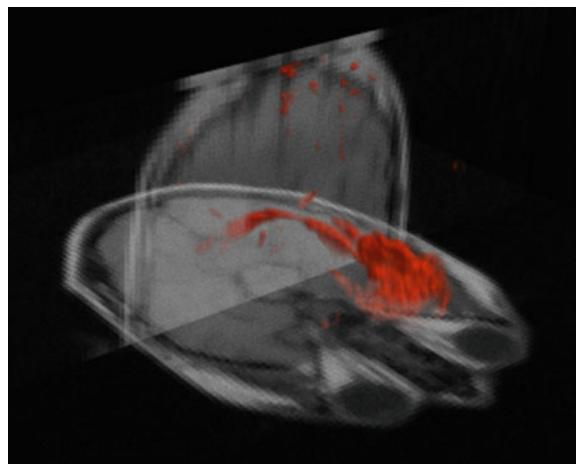
Figure 6 depicts an original abdominal computed tomography (CT) image as well as the color and opacity map. The map shows the bone and soft tissues. Figure 7 shows a rendering image of the slice plane using SPBR. In these figures, the number of particles of (a), (b), and (c) are 16 million, 40 million, and 34 million, respectively. Figure 7 shows that the slice image conforms to the behavior of the opacity map; the bone areas are denoted by yellow pixels. The effect of slice image transparency is shown by the fused visualization (Fig. 8). The transparency implies a sense of depth.

Figure 9 shows magnetic resonance imaging (MRI) images of the head, a T1-weighted image, and a T2-weighted image, as well as their respective color and opacity maps. The maps are generated to show the entire volumetric data. We render Fig. 9a, which is combined with the mapping table (b) and Fig. 9c, which is combined with the mapping table (d). Figure 10 shows the rendering image of the slice plane and volume using SPBR. This slice image, which is the fused axial and



**Fig. 10** Results of the slice image and volume rendering image. **a** Two slice images created from the T1-weighted image. **b** Volume rendering image created from the T2-weighted image

**Fig. 11** Result of fused visualization



coronal plane, is created from Fig. 9a, b; the volume rendering image is created from Fig. 9c, d. In these figures, the number of particles of (a) and (b) are 18 million (including the two slice images) and 6 million, respectively. Figure 10a consists of the two slice images, which are the axial plane and coronal plane. Figure 11 shows the fused visualization for the slice and volume rendering. It is possible to confirm the position and size of the affected region based on the anatomical structure.

## 4 Conclusion

In this paper, we proposed a technique for rendering a non-uniform transparency slice image with a color and opacity map. Using the proposed method, a non-uniform transparent slice is obtained by calculating the number of particles at each cell. For a case in which the area of the particle is greater than the area of the cell—for ex high-resolution volumetric data—the number of particles at each cell cannot be accurately determined. For the latter case, we showed that a subdomain with an area of approximately ten times the area of the particle would suffice. Moreover, we presented fusion images using particle-based rendering, such as volume–slice rendering, based on medical data. The transparency implies a sense of depth. Through the proposed method, the position and size of the affected region can be confirmed based on the anatomical structure.

## References

1. Drebin, R., Carpenter, L., Hanrahan, P.: Volume rendering. In: Proceeding of SIG-GRAPH'88, pp. 29–37 (1988)
2. Levoy, M.: Display of surface from volume data. *IEEE Comput. Graph. Appl.* **8**(3), 29–37 (1988)
3. Carpenter, L.: The a-buffer, an antialiased hidden surface method. *Comput. Graph.* **34**, 34–42 (2010)
4. Everitt, C.: Interactive order-independent transparency. Technical Report NVIDIA Corporation (2001)
5. Koyamada, K., Sakamoto, N., Tanaka, S.: A particle modeling for rendering irregular volumes. In: Proceedings of the International Conference on Computer Modeling and Simulation (UKSIM 2008), Cambridge, England, pp. 372–377 (2008)
6. Sakamoto, N., Kawamura, T., Koyamada, K.: Improvement of particle-based volume rendering for visualizing irregular volume data sets. *Comput. Graph.* **34**(1), 34–42 (2010)
7. Tanaka, S., Hasegawa, K., Shimokubo, Y., Kaneko, T., Kawamura, T., Nakata, S., Ojima, S., Sakamoto, N., Tanaka, H.T., Koyamada, K.: Particle-based transparent rendering of implicit surfaces and its application to fused visualization, EuroVis 2012, Vienna, Austria, June 5–8, 2012
8. Satoshi, T., Akio, M., Satoru, N., Yasushi, F., Hiroaki, Y.: Sampling implicit surfaces based on stochastic differential equations with converging constraint. *Comput. Graph.* **24**(3), 419–431 (2000)
9. Takayama K., Okabe M., Ijiri T., Igarashi T.: Lapped solid textures: filling a model with anisotropic textures. *ACM Trans. Graph. (TOG)* **27**(3) (2008)
10. Hasegawa, K., Hachimura, K., Tanaka, S.: 3D Fused Visualization Based on Particles-Based Rendering with Opacity Using Volume Texture, AsiaSim2013, Singapore, pp. 160–166 (2013)

**Part V**

**Healthcare Support System**

# Automated Diagnosis of Parkinsonian Syndromes by Deep Sparse Filtering-Based Features

Andrés Ortiz, Francisco J. Martínez-Murcia, María J. García-Tarifa,  
Francisco Lozano, Juan M. Górriz and Javier Ramírez

**Abstract** Parkinsonian Syndrome (PS) or Parkinsonism is the second most common neurodegenerative disorder in the elderly. Currently there is no cure for PS, and it has important socio-economic implications due to the fact that PS progressively disables people in their ordinary daily tasks. However, precise and early diagnosis can definitely help to start the treatment in the early stages of the disease, improving the patient's quality of life. The study of neurodegenerative diseases has been usually addressed by visual inspection and semi-quantitative analysis of medical imaging, which results in subjective outcomes. However, recent advances in statistical signal processing and machine learning techniques provide a new way to explore medical images yielding to an objective analysis, dealing with the Computer Aided Diagnosis (CAD) paradigm. In this work, we propose a method that selects the most discriminative regions on  $^{123}\text{I}$ -FP-CIT SPECT (DaTSCAN) images and learns features using deep-learning techniques. The proposed system has been tested using images from the Parkinson Progression Markers Initiative (PPMI), obtaining accuracy values up to 95 %, showing its robustness for PS pattern detection and outperforming the baseline Voxels-as-Features (VAF) approach, used as an approximation of the visual analysis.

## 1 Introduction

Parkinsonian Syndrome (PS) is the second most common neurodegenerative disorder after Alzheimer's disease [5]. It is characterized by a range of motor symptoms, such as bradykinesia, rigidity or tremor, as well as nonmotor symptoms, e.g. fatigue

---

A. Ortiz (✉) · M.J. García-Tarifa · F. Lozano  
Communications Engineering Department, University of Málaga,  
29004 Málaga, Spain  
e-mail: aortiz@ic.uma.es

F.J. Martínez-Murcia · J.M. Górriz · J. Ramírez  
Department of Signal Theory, Communications and Networking,  
University of Granada, 18060 Granada, Spain

and depression. There are a number of entities that together make the PS, although the most frequent cause is the well known Parkinson's Disease (PD). Current diagnosis of PD is mainly based on observable symptoms, along with the administration of medication, such as levodopa, to confirm or refute the diagnosis [15].

A number of neuroimage modalities have been proposed for its use in the differential diagnosis of idiopathic PD, however the existing guidelines allow its use only as a supplementary tool [15]. The image modalities range from structural Magnetic Resonance Imaging (MRI) to functional types, such as SPECT. A number of molecular imaging techniques for presynaptic dopamine activity have been developed to differentiate between parkinsonian syndromes (PD, Multiple System Atrophy -MSA- or Progressive Supranuclear Palsy -PSP-) and other pathologies essential tremors or control. Among them,  $^{123}\text{I}$ -FP-CIT SPECT (also known as its tradename, DaTSCAN) is perhaps the most used radiotracer in clinical practice. It binds to the Dopamine Transporters (DAT) in the caudate and putamen, and reveals normal levels in healthy participants (CTL), patients with essential tremor or drug-induced parkinsonism, and reduced levels in patients with PD, MSA or PSP [14].

Lately, a substantial effort has been put into automatically characterizing the different patterns associated with neurodegenerative diseases, which leaded to the Computed Aided Diagnosis (CAD) paradigm. In the case of DaTSCAN, several works pointed out the advantages of using multivariate decomposition algorithms such as Principal Component Analysis (PCA) [19], Partial Least Squares (PLS) [17] or Independent Component Analysis (ICA) [7, 11], as well as other parameters such as texture features [12]. That way, new feature extraction techniques can be successfully applied to the problem, potentially leading to better accuracy in the diagnosis task.

Deep learning techniques have been successfully applied to different problems involving feature extraction [8, 16]. They are inspired by how the brain works, by establishing a series of layers of deep neural networks where each layer is supposed to extract higher-level features. Sparse Filtering [13] is a new technique for feature extraction based on the optimization of the sparsity of  $\ell_2$ -normalized features that can be used to train deep learning architectures using the greedy layerwise approach [8]. Although the term *deep* is usually linked to networks composed of more than three layers, the use of two layers of sparse filters has demonstrated to provide meaningful and discriminative enough features with the experiments performed. In this work, we propose an algorithm to automatically diagnose PD using deep sparse filtering-based features. First, we proceed by preprocessing the images and selecting a sub-volume that contains the regions of interest (ROIs). After the preprocessing, some deep learning techniques, including sparse filtering and unsupervised feature learning, are applied to the images to obtain a suitable feature vector. Finally, the discriminative power of these features in the PD vs CTL case is evaluated by means of a linear Support Vector Classifier (SVC) [20].

This article is organised as follows. In Sect. 2, a detailed report of the methodology used in this article is presented, along with a description of the database. Later, in Sect. 3 the results of applying our methodology are presented and deeply analysed. Finally, in Sect. 4, some conclusions are drawn.

## 2 Methodology and Materials

### 2.1 Database

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database ([www.ppmi-info.org/data](http://www.ppmi-info.org/data)). For up-to-date information on the study, visit [www.ppmi-info.org](http://www.ppmi-info.org).

Images in this database were acquired 4, 5h after the injection of between 111 and 185 MBq of DaTSCAN. Raw projection data are acquired into a  $128 \times 128$  matrix stepping each 3 degrees for a total of 120 projection into two 20 % symmetric photopeak windows centered on 159 and 122 KeV with a total scan duration of approximately 30–45 min. The images of the subject's data are reconstructed and attenuation corrected, implementing either filtered back-projection or an iterative reconstruction algorithm using standardized approaches [18].

### 2.2 Image Preprocessing

The preprocessing stage consist of two steps: spatial and intensity normalization. As PPMI images are spatially normalized as described in [10], we only perform intensity normalization to homogenize the intensity scale throughout the database. Later, a binary mask is applied to select a box-shaped volume that encloses the Regions of Interest (ROIs) associated to PD.

### 2.3 Intensity Normalization

As commented above, intensity normalization is required to be able to compare the uptake value in areas of specific activity (related to dopaminergic transporters) and areas of non-specific activity (vascular activity) between subjects. The method used in this work is similar to the one described in [12], consisting on computing an intrinsic parameter  $I_p$  from the image, and the estimation of the binding activity as

$$t' = t/I_p \quad (1)$$

where  $t$  denotes the spatially normalized image and  $t'$  the spatial and intensity normalized image.

The normalization value  $I_p$  is obtained by means of the *Integral Normalization* method, which approximates the value  $I_p = \int t$  as the sum of intensity values of the image, giving an integral value of the intensity.

## 2.4 Masking

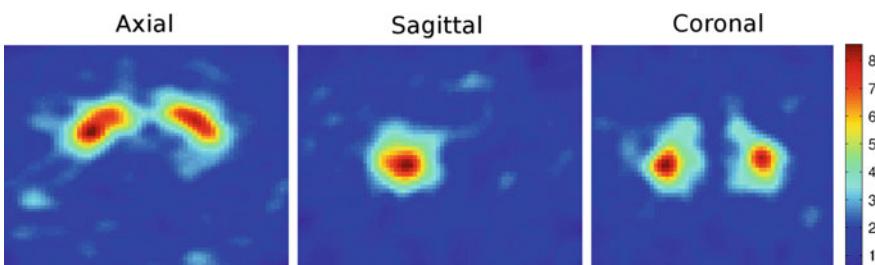
Once the images have been normalized, a sub-volume of the whole images can be selected to extract the relevant regions. This masking procedure has a double purpose. On the one hand, it avoids driving the subsequent stages (such as feature extraction algorithms) with non-relevant voxels (since discriminative information in DaTSCAN images is focused in the striatum). On the other hand, the data dimensionality is considerably reduced, resulting in less computational burden.

To perform this selection, we use the algorithm described in [12]. The algorithm first averages all the images in the database to create a mean image  $I_{mean}$ . Later, a threshold  $I_{th}$  is set depending on the characteristics of the images. Due to the nature of the DaTSCAN images, the higher intensity as well as the major differences between affected patients and controls are located at the striatum. Therefore, the threshold  $I_{th}$  can be defined in terms of the intensity values of the images. However,  $I_{th}$  has to be properly selected as voxels out of the selected region will not be used in further operations. Since the ROIs in this problem are located in the highest activation regions, it is possible to determine the threshold as the central value of intensity in the whole image, and computed as

$$I_{th} = \frac{1}{2}(max(I_{mean}) - min(I_{mean})) + min(I_{mean}) \quad (2)$$

where  $I_{mean}$  is the mean of all the images in the training set.

Finally, after  $I_{th}$  is set, we create a box-shaped mask that contains only brain voxels with an intensity level higher than a computed threshold  $I_{th}$ . This way, we select the ROIs as well as their surrounding areas in each image for further processing. Figure 1 shows an example of the extracted volume for an image from the PPMI database using the method explained above.



**Fig. 1** Example of extracted volume containing regions of interest related to dopamine activity for an image from the PPMI database

## 2.5 Sparse Filtering

Unsupervised feature extraction deals with computing a number of predictors that accurately represent the classes present in the data manifold, without the use class label information. This way, algorithms such as Principal Component analysis (PCA) [1, 3, 9], Independent Component Analysis (ICA) [2, 7, 11] or Gaussian Mixture Models (GMM) [6], aim to extract features by building a model based on the underlying data distribution. By contrast, Sparse Filtering [13] does not explicitly attempt to construct a model based on the data distribution but it is based on the optimization of the sparsity of  $\ell_2$ -normalized features. Sparse filtering method learn features in an unsupervised way while dealing with (1) population sparsity (each sample should be represented by a sparse vector, i.e. with only few non-zero components), (2) Lifetime sparsity (each feature should be non-zero only for a few samples) and (3) High dispersal (the activity of the features has to be uniformly distributed). Let  $f_j^{(i)}$  represent the  $j$ -th feature value for the  $i$ -th sample in a manifold composed of  $M$  samples. The method described in [13] consists on first normalizing each feature by its  $\ell_2$ -norm across all samples and the normalizing all feature for all samples, as shown in Eqs. 3 and 4, respectively.

$$\tilde{f}_j = \frac{f_j}{\|f_j\|_2} \quad (3)$$

$$\hat{f}_i = \frac{\tilde{f}_i}{\|\tilde{f}_i\|_2} \quad (4)$$

Subsequently, the sparsity of the normalized features is minimized by means of minimizing the objective function

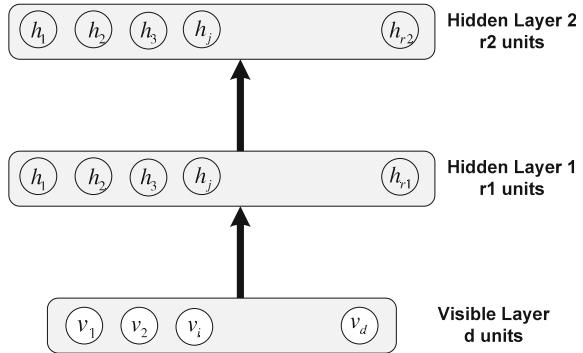
$$\sum_{i=1}^M \|\hat{f}^{(i)}\|_1 = \sum_{i=1}^M \left\| \frac{\hat{f}^{(i)}}{\|\hat{f}^{(i)}\|_2} \right\|_1 \quad (5)$$

using the Limited Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method, where the term  $\|\hat{f}^{(i)}\|_1$  represents the population sparsity of the features on the  $i$ th sample.

## 2.6 Deep Neural Networks and Sparse Filtering

Sparse filtering method explained above, learn a number of feature extractors in an unsupervised way. Thus, these features can be used as weights of the hidden layers in a feedforward neural network, even using non-linear activation functions that will determine the activated neurons (features) for new samples. Consequently, a deep neural structure can be unsupervisedly trained using the greedy layerwise approach

**Fig. 2** Deep neural structure used compute features



[4], which consist on using the weights of a single layer as inputs to train the another (higher) layer corresponding to more abstract features.

In this work we train a three layer feedforward neural network as shown in Fig. 2

## 2.7 Unsupervised Feature Learning

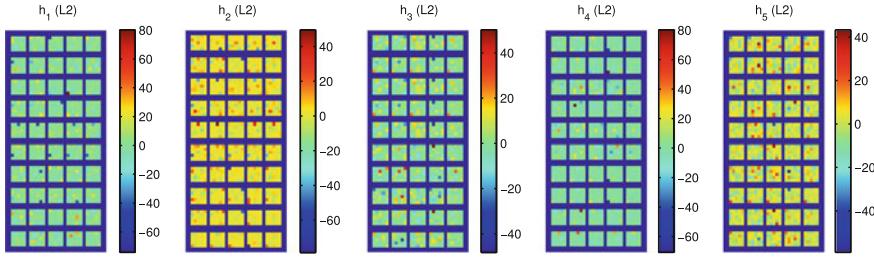
Feature extraction is addressed in this work using a three layer neural network which is trained in an unsupervised way by means of sparse filtering. Thus, raw voxels extracted from the volume containing the PD-related ROI (as explained in Sect. 2.4), are used to train the first layer. Moreover, these ROIs are split into non-overlapped 3D windows containing  $d$  voxels each. Thus, each non-overlapped window is used as a sample in the training process using the  $d$  voxels composing each window to feed the  $d$  neurons composing the input layer.

Unsupervised learning is performed by Sparse Filtering, which computes the weights connecting input units to the first layer units and consequently. Then, the input is feedforwarded to the first layer using the activation function described in [13]

$$f_j^{(i)} = \sqrt{\epsilon + (\omega_j^T x^{(i)})^2} \quad (6)$$

Subsequently, the  $f$  is normalized first by rows (features) and then by columns (samples) as described in Sect. 2.5. Once the activations of the first layer are computed, these are used as inputs to train the second layer, using the greedy layerwise approach [8].

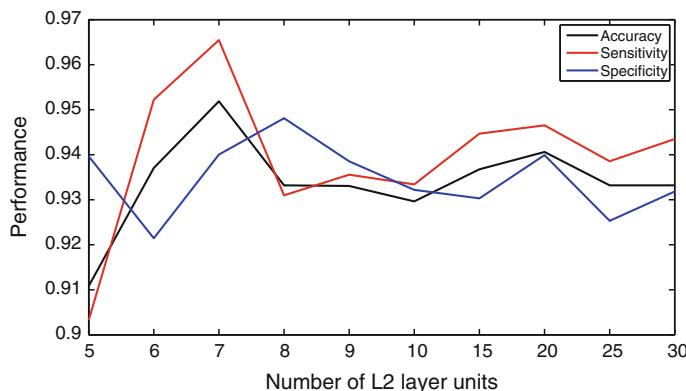
As a result of the training process described above, features at different abstraction levels are computed. Thus, Fig. 3 shows the selected filters from layer 1, corresponding to the strongest connections to layer 2 units (i.e. filters shown correspond to units with highest absolute weights).



**Fig. 3** Ten best Layer 1 units ( $h_i^1$ ) selected using the maximum absolute coefficient of the different units  $h_i^2$  in layer 2

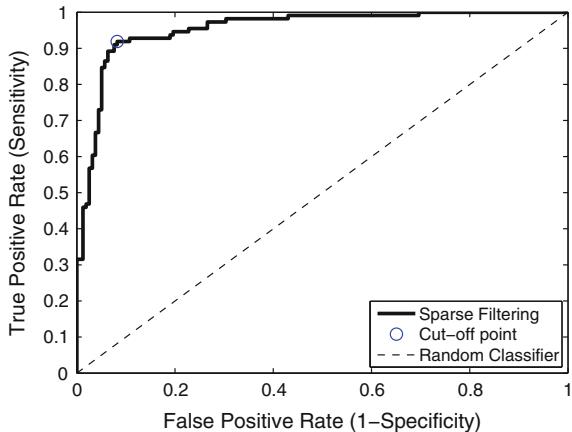
### 3 Results and Discussion

Once the neural structure shown in Fig. 2 is trained by sparse filtering, it is used to compute features from new images. This is addressed by presenting the 3D patches extracted from the ROIs related to dopamine activity (as indicated in Sect. 2.4) to the visible layer of the network. Then, these voxel values are feed-forwarded to the first layer and eventually to the second layer. The activation level of layer 2 units are then used as features that are concatenated to describe the former image. The experimental setup used in this work includes cubic 3D patches of  $5 \times 5 \times 5$  voxels and 80 units at L1. As a consequence, the visible layer is composed of 125 units. The discriminative capabilities of the computed features has been assessed by feeding a linear Support Vector Classifier (SVC) with the concatenation of all the features extracted from the ROIs of an image. In this work, two experiments have been conducted. Firstly, we vary the number of L2 units, which corresponds to the number of



**Fig. 4** Classification performance. Accuracy, sensitivity and specificity are shown for different number of L2 units

**Fig. 5** ROC curve of the classification method.  
AUC = 0.95



features extracted per 3D patch. This resulted in the graph shown in Fig. 4. Specifically, accuracy, sensitivity and specificity are computed for different number of L2 units.

The results of this experiments revealed the best performance is achieved for 7 units at L2, where accuracy values are over 0.95. One can see that our system clearly outperforms the baseline Voxels-As-Features (VAF) approach, which has an accuracy of 0.84 (and a sensitivity and specificity of 0.807 and 0.862 respectively).

In the second experiment, the ROC curve was computed using 7 units at L2 with all the remaining parameters fixed. The SVC distance to hyperplane was varied to compute the ROC curve shown in Fig. 5, obtaining an Area Under the Curve (AUC) of 0.95. Both experiments prove the ability of the sparse filtering strategy to identify and extract relevant PD features in DaTSCAN images, and revealing as an accurate tool to help in a medical environment.

## 4 Conclusions and Future Work

This paper proposes the use of sparse filtering to extract features from the regions of interest in order to detect the Parkinson's disease. Specifically, voxels from regions related to dopaminergic activity are used. Thus, Sparse filtering technique is used to unsupervisedly train a two-layers network, dealing with learning features from the image data without using label information. Subsequently the trained network is used to compute features from new images.

Classification experiments conducted to explore the influence of the parameters involved in the feature extraction process revealed the optimal number of per-window features to obtain the best accuracy value (0.95). In addition, this number of features provided sensitivity and specificity values of 0.96 and 0.94, respectively along with an AUC value derived from the ROC curve of 0.95. As a conclusion, sparse filtering provides discriminative-enough features to detect the Parkinson's disease from

image data. We also plan to explore the effect of using a deeper network and to analyse the features extracted at each layer, which should be more abstract and may reveal more specific, disease-related structures within the ROIs.

**Acknowledgments** This work was partly supported by the MINECO under the TEC2015-64718-R and PSI2015-65848-R projects and the Consejería de Economía, Innovación, Ciencia y Empleo (Junta de Andalucía, Spain) under the P11-TIC-7103 Excellence Project. PPMI—a public-private partnership—is funded by the Michael J. Fox Foundation for Parkinson’s Research and funding partners, including Abbvie, Avid Radiopharmaceuticals, Biogen, Bristol-Myers Squibb, Covance, GE Healthcare, Genetech, GlaxoSmithKline, Eli Lilly and Co., Lundbeck, Merck, MSD Meso Scale Discovery, Pfizer, Piramal, Roche, Servier and UCB.

## References

1. Abdi, H., Williams, L.J.: Principal Component Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, vol. 2, pp. 433–459 (2010)
2. Álvarez, I., Górriz, J.M., Ramírez, J., Salas-González, D., López, M., Segovia, F., Padilla, P., García, C.: Projecting independent components of spect images for computer aided diagnosis of Alzheimer’s disease. Pattern Recogn. Lett. **31**(11), 1342–1347 (2010)
3. Álvarez, I., Górriz, J.M., Ramírez, J., Salas-Gonzalez, D., Lopez, M.M., Segovia, F., Chaves, R., Gómez-Rio, M., Garcia-Puntonet, C.: 18f-fdg pet imaging analysis for computer aided Alzheimer’s diagnosis. Inf. Sci. **184**(4), 903–916 (2011)
4. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., Montréal, Université, Québec, Montréal: Greedy Layer-Wise Training of Deep Networks. MIT Press, In NIPS (2007)
5. de Lau, L.M., Breteler, M.M.: Epidemiology of parkinson’s disease. Lancet Neurol **5**(6), 525–535 (2006)
6. Górriz, J.M., Segovia, F., Ramírez, J., Lassl, A., Salas-González, D.: Gmm based spect image classification for the diagnosis of Alzheimer’s disease. Appl. Soft Comput. **11**, 2313–2325 (2011)
7. Graña, M.: Towards relevance dendritic computing. In: Proceedings of the Nature and Biologically Inspired Computing (NaBIC), pp. 588–593 (2011)
8. Hinton, G.E., Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. Neural Comput. **18**(7), 1527–1554 (2006)
9. López, M., Ramírez, J., Górriz, J.M., Álvarez, I., Salas-González, D., Segovia, F., Chaves, R., Padilla, P., Gómez-Río, M.: Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer’s disease. Neurocomputing **74**(8), 1260–1271 (2011)
10. Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburz, K., Flagg, E., Chowdhury, S., Poewe, W., Mollenhauer, B., Sherer, T., Frasier, M., Meunier, C., Rudolph, A., Casaceli, C., Seibyl, J., Mendick, S., Schuff, N., Zhang, Y., Toga, A., Crawford, K., Ansbach, A., de Blasio, P., Piovella, M., Trojanowski, J., Shaw, L., Singleton, A., Hawkins, K., Eberling, J., Russell, D., Leary, L., Factor, S., Sommerfeld, B., Hogarth, P., Pighetti, E., Williams, K., Standaert, D., Guthrie, S., Hauser, R., Delgado, H., Jankovic, J., Hunter, C., Stern, M., Tran, B., Leverenz, J., Baca, M., Frank, S., Thomas, C.A., Richard, I., Deeley, C., Rees, L., Sprenger, F., Lang, E., Shill, H., Obradov, S., Fernandez, H., Winters, A., Berg, D., Gauss, K., Galasko, D., Fontaine, D., Mari, Z., Gerstenhaber, M., Brooks, D., Malloy, S., Barone, P., Longo, K., Comery, T., Ravina, B., Grachev, I., Gallagher, K., Collins, M., Widnell, K.L., Ostrowizki, S., Fontoura, P., La-Roche, F.H., Ho, T., Luthman, J., van der Brug, M., Reith, A.D., Taylor, P.: The Parkinson progression marker initiative (PPMI). Progr. Neurobiol. **95**(4), 629–635 (2011)

11. Martínez-Murcia, F.J., Górriz, J.M., Ramírez, J., Illán, I.A., Ortiz, A.: Automatic detection of parkinsonism using significance measures and component analysis in datscan imaging. *Neurocomputing* **126**, 58–70 (2014)
12. Martínez-Murcia, F.J., Górriz, J.M., Ramírez, J., Moreno-Caballero, M., Gómez-Río, M.: Parametrization of textural patterns in  $^{123}\text{I}$ -ioflupane imaging for the automatic detection of parkinsonism. *Med. Phys.* **41**(1), 012502 (2014)
13. Ngiam, J., Chen, Z., Bhaskar, S.A., Koh, P.W., Ng, A.Y.: Sparse filtering. In: Shawe-Taylor, J., Zemel, R.S., Bartlett, P.L., Pereira, F., Weinberger, K.Q. (eds.) *Advances in Neural Information Processing Systems*, vol. 24, pp. 1125–1133. Curran Associates, Inc. (2011)
14. Ortega Lozano, S.J., Martínez Del Valle Torres, M.D., Ramos Moreno, E., Sanz Viedma, S., Amrani Raissouni, T., Jiménez-Hoyuela, J.M.: Quantitative evaluation of spect with fp-cit. importance of the reference area. *Rev. Esp. Med. Nucl.* **29**(5), 246–250 (2010)
15. Politis, M.: Neuroimaging in parkinson disease: from research setting to clinical practice. *Nat. Rev. Neurol.* **10**(12), 708–722 (2014)
16. Raja, K.B., Raghavendra, R., Krishna Vemuri, V., Busch, C.: Smartphone based visible iris recognition using deep sparse filtering. *Pattern Recogn. Lett.* **57**, 33–42 (2015)
17. Segovia, F., Manuel Górriz, J., Ramírez, J., Chaves, R., Álvarez Illán, I.: Automatic differentiation between controls and parkinson's disease datscan images using a partial least squares scheme and the fisher discriminant ratio. In: *KES*, pp. 2241–2250 (2012)
18. The Parkinson Progression Markers Initiative: PPMI. Imaging Technical Operations Manual, 2nd edn (2010)
19. Towey, D.J., Bain, P.G., Nijran, K.S.: Automatic classification of  $^{123}\text{I}$ -fp-cit (datscan) spect images. *Nucl. Med. Commun.* **32**(8), 699–707 (2011)
20. Vladimir, N.: *Vapnik*. Wiley-Interscience, Statistical Learning Theory (1998)

# Post-stroke Hand Rehabilitation Using a Wearable Robotic Glove

Dorin Popescu, Mircea Ivanescu, Razvan Popescu, Anca Petrisor,  
Livia-Carmen Popescu and Ana-Maria Bumbea

**Abstract** The paper presents the research work done for development of a light-weight and low-cost robotic glove that post-stroke patients can use to recover hand functionality. The work focused on two directions for the robotic glove structure (exoskeleton and wearable soft robotic glove) and on two types of recuperative actions (tele-operation and program based actions). Given the performance tests ran for the robotic gloves, better results were shown with the wearable soft robotic glove that could also be combined with Functional Electrical Stimulation in order to improve the post-stroke hand rehabilitation.

**Keywords** Stroke • Robotic glove • Rehabilitation • Functional electrical stimulation

## 1 Introduction

Health is the most important issue for everybody. The cerebrovascular accident (or stroke) is the second leading cause of death. World Health Organization (WHO), as responsible for international health coordination, monitors and assesses health trends in order to plan and implement health policies. WHO has developed an international stroke surveillance system: the STEPwise approach to stroke surveillance (STEPS-stroke), described on its website (<http://www.who.int/en>).

A cerebrovascular accident (CVA) occurs when the blood supply to the brain is interrupted or reduced, which deprives the brain of oxygen and nutrients and can cause brain cells to die or death in case of a severe stroke.

Stroke patients often lose certain hand functions. It is difficult to open the affected hand to grasp due to increased resistance to passive finger extension and

---

D. Popescu (✉) · M. Ivanescu · R. Popescu · A. Petrisor · L.-C. Popescu  
University of Craiova, Craiova, Romania  
e-mail: dorinp@robotics.ucv.ro

A.-M. Bumbea  
University of Medicine and Pharmacy, Craiova, Romania

weakness in finger extensors [1]. Because of this impairment, post-stroke hand rehabilitation is essential for restoring independent behaviour.

The aim of our work was to design and develop an Intelligent Haptic Robot-Glove (IHRG) for post-stroke hand rehabilitation. Our work focused on developing a lightweight and low-cost robotic glove that patients can use to recover hand functionality.

The rehabilitation aims to help stroke patients to relearn the skills that were lost when they suffered the stroke. Any rehabilitation program means repetitive practice and begins in the hospital after the patient's overall condition has been stabilized (often within 24–48 h after the stroke) and could involve working with physiotherapist for months or years after the stroke.

Maciejasz, Eschweiler, Gerlach-Hahn, Jansen-Troy, and Leonhardt conducted a survey on robotic devices for upper limb rehabilitation, that includes a comprehensive, tabulated comparison of technical solutions implemented in various systems [2]. The control area of these systems is very large and complex and imposes a series of requirements regarding the size, the cosmetic appearance, robustness and embeddable control system.

A review of control strategies for robotic movement training after neurologic injury was presented by Marchal-Crespo and Reinkensmeyer [3]. The control must be simple with actuator able to exert high grasping forces [4]. A control strategy based on neural approach is presented by Rodriguez-Cheu and Casals [5]. A hybrid control technique is treated by Zhao et al. [6]. A digital controller that operates by means the EMG signal is analysed by Lucas et al. [7]. The feedback control by sensors is treated by Scherillo et al. [8], Birglen and Gosselin [9], Krut [10], Birglen and Gosselin [11], Xiujuan and Zhen [12]. A virtual reality system was developed in order to encourage repetitive task practice [13]. Luo et al. [14] developed a training environment that integrates augmented reality (AR) with assistive devices for post-stroke hand rehabilitation.

Matheson and Brooker [15] developed an exoskeleton to provide the actuation for flexion. Pneumatic muscle actuators were used to provide flexion force, and force sensors used to supply control inputs by the user. Hartopanu et al. [16] proposed a solution to combine a robotic glove with Functional Electrical Stimulation to be used as a tool in the rehabilitation process of patients who suffered a stroke.

Neurofeedback can improve rehabilitation when patients get immersive feedback that relates to the activities they imagine or perform. For example, if people imagine grasping an object with their affected hand, then an image of a grasping hand can help users visualize their activity. In the last years, a totally novel application for motor imagery (MI)—based Brain-Computer Interface (BCI) has gained attention by inducing neural plasticity and thus serve as an important tool to enhance motor rehabilitation for stroke patients [17]. Furthermore, immersive BCI stroke rehabilitation is an ongoing research effort in numerous American and European research projects (<http://www.gtec.at/Research/Projects>). The BCI system can also be connected to exoskeletons or rehabilitation devices.

## 2 Post-stroke Hand Rehabilitation Therapy

Because a part of the brain is damaged when a stroke occurs, the rehabilitation therapy will help the patient regain his/her lost hand functions in a non-damaged part of his/her brain. Functional Electrical Stimulation—FES could help the rehabilitation therapy through reducing spasticity and enhancing muscle control [18, 19].

Usually, post-stroke hand rehabilitation includes passive movements or exercises (movements done with the help of a physiotherapist) and more active exercises patients do with little or no assistance. Stroke rehabilitation can be exhausting and daunting.

Post-stroke rehabilitation has a lot of new approaches, that include new kinetic techniques, new electrical stimulation techniques, and in the last years, robotics methods.

The research team received from doctors a challenge to develop a robotic device to help the patients with stroke to recover the hand movement through an Intelligent Haptic Robot-Glove (IHRG).

The following set of guidelines were taken into account when designing and developing the robotic glove:

- natural/intuitive use and easy to wear.
- lightweight.
- not restricting the natural human kinematics or range motion.
- compliance and stability.
- sufficient adaptability to individual differences in patients' anthropometric dimensions (without mechanical regulation or tunings).
- reduced system costs.
- easy maintenance.
- high power-to-weight density and reduced energy consumption.

Given these requirements, our goal was to create a wearable robotic glove that assists the patient during hand movement.

Post-stroke hand rehabilitation is time dependent. Good results of post-stroke rehabilitation therapy could be obtained in the first 6 months after the stroke occurred.

The muscle tonus is very weak in the first post-stroke stage. In time the muscle tonus increases, but not in a healthy way. It could be very spastic, which can affect the future movements. To avoid this outcome, in the first stage, if the muscle palsy lasts for too long time, it is indicated to apply electrical stimulation to the flexor muscles, to develop the spasticity in a gentle manner. During the second stage, when the spasticity is created, the electrical stimulation has to be applied to the extensor muscles of hand and fingers to maintain muscle balance.

Nowadays we have many types of electrical stimulation which include the square waveform signals or magnetic field, applied to extensor muscles. The signal parameters are modified in every stage and are determined by the spasticity.

At the beginning, electrical stimulation with the square waveform signals is applied in short pulses, which determines an isotonic muscle contraction,

followed by a double period of muscle relaxation. We can use this method for a few minutes, but no more than 10 min or we risk to develop muscle fatigue and improper hand movement.

Continuing the electrical stimulation technique, the square waveform signal could be applied in long pulses followed by double or triple periods of muscle relaxation. The time of this application is determinated by muscle fatigue and false movement. The session of electrical stimulation can be repeated twice a day, if the muscle fatigue has not appeared. This technique is recommended to be used for a long period such as weeks or months, until the patient develops correct active hand movement.

The patient must watch the movement and try to move the hand by him/her self while also imagining his/her hand movement (neurofeedback).

A new technique for post-stroke hand rehabilitation is the robotic approach, by using a robotic glove. In order to design and develop such a robotic glove we went through a series of research phases (survey, video motion analysis, virtual design, models design, mechanical design, actuation system design, control system design, implementation, testing, refinement).

### 3 Robotic Glove Design

A survey on the current state of research on the mechanisms and achievements that model or substitute the human hand was conducted. Thus we could identify two categories of systems that model or substitute the human hand:

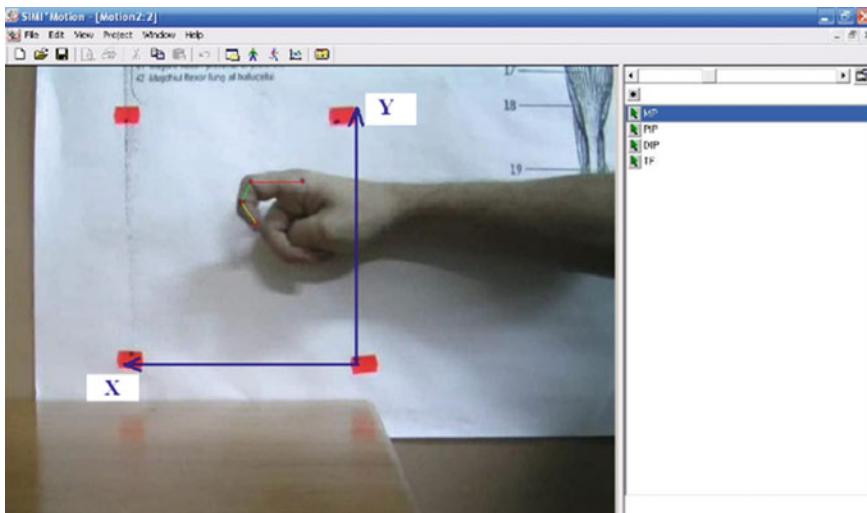
- grasping anthropomorphic mechanical systems for prosthetics;
- grasping anthropomorphic mechanical systems for robots.

We decided to research on two directions for the robotic glove structure: exoskeleton and soft robotic glove. The soft robotic glove is a device that uses textiles to interface with the body in parallel with the muscles, using the bone structure, to support fingers' motion. We focused on designing a soft robotic glove that could use actuators and sensors that don't restrict hand movement.

Our research also addressed the study and design of an exoskeleton model for a robotic glove (as an assistance hand) and its associated actuation and control systems. This robotic glove must be attached to the human hand and allow the hand and fingers movements. Based on these requirements (the movement in different planes adapted to the patient's hand, the possibility to touch and grasp), specific biomechanical design of the components was done.

The robotic glove has to provide three features for grasping force:

1. The robotic glove should not disturb human finger movement.
2. The robotic glove has to allow a grasping force proportional to the human grasping force.
3. Robotic glove has to allow a variable compliance as the human finger so that the dexterity and stability of the grasping is preserved [20].



**Fig. 1** Index finger motion analysis using SimiMotion software

Other research phase consisted of determining the variation laws of kinematic parameters of human hand movement based on studying its movement using a data acquisition system (video motion analysis) called SimiMotion ([www.simi.com](http://www.simi.com)) (Fig. 1). In order to implement good movement for the robotic glove fingers, we analysed using image processing the motion of the human hand fingers and then we developed a tele-operated robotic hand accordingly (Fig. 2).

Based on the above analysis two types of robotic gloves were designed, developed and tested, including mechanical structures and actuation systems.

**Fig. 2** Tele-operated robotic hand



## 4 Results

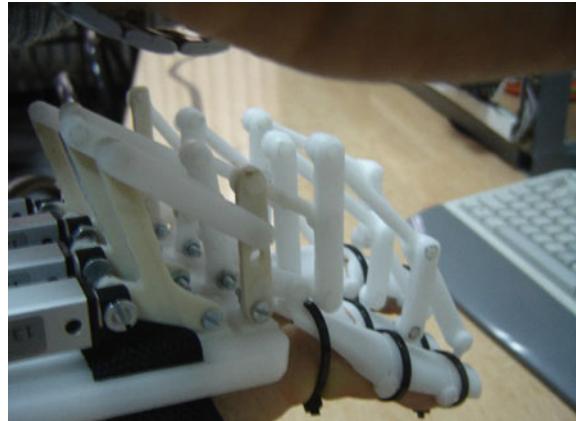
We focused on creating an assistive device (robotic glove) that provides post-stroke hand rehabilitation and showed that this system can substantially maintain normal biomechanics.

The robotic glove consists of three parts: mechanical exoskeleton or soft robotic glove, actuation system, and embedded control with minimum number of sensors.

To address the rehabilitation issues, we developed a lightweight exoskeleton that allows for basic motion using a natural sequence of muscle activation.

The exoskeleton system has a kinematic structure with five articulated fingers. Each finger is composed of three phalanges in order to have an anthropomorphic contact with the patient's hand. In order to get an effective rehabilitation effect, the mechanical structure must allow the finger to fulfill the same motions of a healthy finger.

**Fig. 3** Exoskeleton structure



**Fig. 4** Wearable soft robotic glove with Bowden tendons



**Fig. 5** Wearable soft robotic glove with flexible tendons



We designed and tested different mechanical solutions, different motion transmission solutions and actuation systems (electric, pneumatic and SMA actuation systems).

Two structures were designed and developed:

- exoskeleton—the structure with phalanges (Fig. 3).
- soft robotic glove—the glove type structure (Figs. 4 and 5).

We designed, developed and tested two solutions in order to drive the movement from the actuation system to the robotic glove:

- through tendons (Bowden type (Fig. 4) or flexible tendons (Fig. 5) for the glove type structure).
- through bar mechanism for the structure with phalanges.

An embedded control system based on an Arduino Mega2560 board was developed and tested.

Two types of recuperative actions were implemented:

- tele-operation using a glove with flex sensors.
- program based actions; the doctor can choose a specific recuperative program or can create new ones.

Comparison of the actuation systems shows:

- Pneumatics (cylinders or muscles): difficult to control; voluminous; expensive.
- SMA: difficult to develop; small force; small stroke and complexity of position control.
- Electric actuators:
  - rotary actuators: problems to convert rotary movement in linear movement (need of additional elements).
  - linear actuators: easy to drive the linear movement; no difficult to control (Fig. 6).

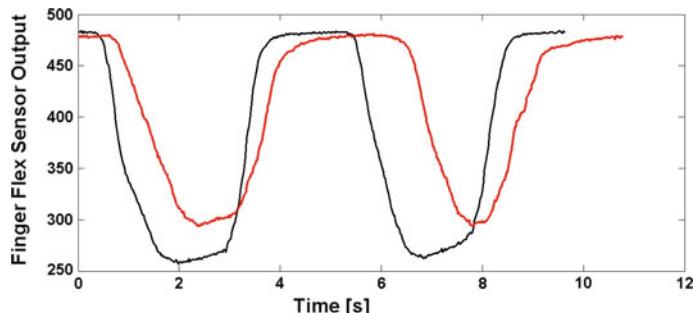


**Fig. 6** Wearable soft robotic gloves in action



**Fig. 7** Robotic glove and FES. Tele-operation glove with flex sensors

Figure 7 presents a variant of the hybrid FES—flexible robotic glove system, which was tested in our laboratory. The tests were made on normal human subjects (members of the research team) who attempted to mimic the behavior of a post-stroke patient, described behavior by doctors and patients.



**Fig. 8** Evolutions in time of the output of the index finger flex sensors

Figure 8 presents the evolutions in time of the output of the index finger flex sensors, one is on tele-operation glove (black line) and the other is mounted on the robotic glove worn by patient (red line). There is a small delay and the flexion of the robotic glove is smaller than of tele-operation glove.

## 5 Conclusions

The goal of our work was to create a robotic glove for patients as well as a tele-operation wearable glove integrated with flex sensors used to control the finger movements of the robotic glove.

Wearable soft robotic gloves seem to be the better solution for human hand rehabilitation as these gloves are lightweight, portable, and compliant wearable systems that can still be further refined.

The paper presents the hardware and software development that has been designed and implemented for the hand rehabilitation robotic glove. The practical tests showed the functionality and performances of the robotic gloves in common operating conditions.

The soft robotic gloves architectures will be more frequently used in assistive and rehabilitation robotics, due to their simplicity. Many development challenges and plenty of research remain to be done in actuator development, textile innovation, soft sensor development, human-machine interface (control), and biomechanics.

The architectures of the exoskeleton robotic glove require for the robotic glove and human joints to be accurately aligned, as misalignments would result in kinematic incompatibilities impeding the movement of the fingers or causing physical discomfort.

We consider this robotic glove not only shortens rehabilitation time post-stroke, but also brings a higher level of recovery than is achievable with the current physiotherapy methods.

The robotic glove is easy to use, and the setting time is short (less than 10 min), as doctors want. The next phase is clinical use followed by clinical research.

Our robotic glove can be combined with FES and neurofeedback to improve the rehabilitation, by making a better hand functionality recovery in a shorter time.

**Acknowledgment** This work is supported by PNCDI-II-PCCA 150/2012 grant of the Romanian Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI).

## References

1. Brokaw, E.B., Black, I., Holley, R., Lum, P.: Hand spring operated movement enhancer (HandSOME): a portable passive hand exoskeleton for stroke rehabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **19**(4), 391–398 (2011)

2. Maciejasz, P., Eschweiler, J., Gerlach-Hahn, K., Jansen-Troy, A., Leonhardt, S.: A survey on robotic devices for upper limb rehabilitation. *J. NeuroEng. Rehabil.* **11**, 3 (2014)
3. Marchal-Crespo, L., Reinkensmeyer, D.J.: Review of control strategies for robotic movement training after neurologic injury. *J. NeuroEng. Rehabil.* **6**, 20 (2009)
4. Wege A., Hommel, G.: Development and control of a hand exoskeleton for rehabilitation of hand injuries. In: Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3046–3051. Edmonton, Canada (2005)
5. Rodriguez-Cheu, L.E., Casals, A.: Sensing and control of a prosthetic hand with myoelectric feedback. In: Proceedings IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, BioRob, pp. 607–612. Pisa, Italy (2006)
6. Zhao, J., Xie, Z., Jiang, L., Cai, H., Liu, H., Hirzinger, G.: A five fingered underactuated prosthetic hand control scheme. In: Proceedings IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, BioRob, pp. 995–1000. Pisa, Italy (2006)
7. Lucas, L., DiCicco, M., Matsuoka, Y.: An EMG-controlled hand exoskeleton for natural pinching. *J. Robot. Mechatron.* **16**(5), 1–9 (2004)
8. Scherillo, P., Siciliano, B., Zollo, L., Carrozza, M.C., Guglielmelli, E., Dario, P.: Parallel force/position control of a novel biomechatronic hand prosthesis. In: Proceedings IEEE/ASME International Conference on Advanced Intelligent Mechatronics, vol. 2, pp. 920–925 (2003)
9. Birglen, L., Gosselin, C.M.: Fuzzy enhanced control of an underactuated finger using tactile and position sensors. *IEEE Trans. Industr. Electron.* **44**(4), 732–738 (2002)
10. Krut, S.: A Force-isotropic underactuated finger. In: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2325–2331. Barcelona, Spain (2005)
11. Birglen, L., Gosselin, C.M.: Geometric design of three-phalanx underactuated fingers. *J. Mech. Des.* **128**(2), 356–364 (2005)
12. Xiujuan, D., Zhen, L.: Underactuated robot dynamic modelling and control based on embedding model. In: Proceedings 12th IFToMM World Congress, Besançon, France (2007)
13. Tsoupikova, D., Stoykov, S.N., Corrigan, M., Thielbar, K., Vick, R., Li, Y., Triandafilou, K., Preuss, F., Kamper, D.: Virtual immersion for post-stroke hand rehabilitation therapy. *Ann. Biomed. Eng.* **43**(2), 467–477 (2015)
14. Luo, X., Kline, T., Fischer, H.C., Stubblefield, K.A., Kenyon, R.V., Kamper, D.G.: Integration of augmented reality and assistive devices for post-stroke hand opening rehabilitation. In: Proceedings of the 27th Annual Conference IEEE on Engineering in Medicine and Biology, vol. 7, pp. 6855–6888. Shanghai, China (2005)
15. Matheson, E., Brooker, G.: Assistive rehabilitation robotic glove. In: Proceedings of Australasian Conference on Robotics and Automation, pp. 1–10. Monash University, Melbourne Australia (2011)
16. Hartopanu, S., Poboroniuc, M.S., Serea, F., Irimia, D., Livint, Gh.: New Issues on FES and robotic glove device to improve the hand rehabilitation in stroke patients. In: Proceedings of the 6th International Conference on Modern Power Systems, pp. 123–127. Cluj-Napoca, Romania (2015)
17. Coffey, A.L., Leamy, D.J., Ward, T.E.: A novel BCI-controlled pneumatic glove system for home-based neurorehabilitation. In: Proceedings Conference IEEE Engineering in Medicine and Biology Society, pp. 3622–3625 (2014)
18. Irimia, D.C., Poboroniuc, M.S., Pasol, I., Ortner, R.: Correlations between muscular contraction type and muscle electrical activity. In: Proceedings of the 15th International Scientific Conference on Electric Power Engineering (EPE), pp. 488–491. Iasi (2014)
19. Hartopanu, S., Poboroniuc, M.S., Serea, F., Livint, Gh.: Towards human arm rehabilitation in stroke patients by means of a hybrid FES-robotic glove. In: Proceedings of the 15th International Scientific Conference on Electric Power Engineering (EPE), pp. 148–152. Iasi (2014)
20. Hasegawa, Y., Mikami, Y., Watanabe, K., Sankai, Y.: Five-fingered assistive hand with mechanical compliance of human finger. In: Proceedings IEEE International Conference on Robotics and Automation (ICRA), pp. 718–724. Pasadena, USA (2008)

# An mHealth Application for a Personalized Monitoring of One's Own Wellness: Design and Development

Manolo Forastiere, Giuseppe De Pietro and Giovanna Sannino

**Abstract** Behavior and lifestyle are the key determinants of health, disease, disability and premature mortality. There is important evidence that demonstrates that unhealthy behaviors increase the risk of the onset of many diseases and therefore could be considered among the causes of the disease itself. The ambition of this app is to provide people with something similar to a personal trainer, an application that, after collecting a range of information on the individual, is able to classify her/him based on her/his individual characteristics (physical parameters and lifestyle) and then to propose specific recommendations to improve her/his well-being. By monitoring the evolution over time of these individual characteristics, the application can also give feedback on the effectiveness of the measures and therefore provide positive stimuli to motivate the user to continue the path taken.

**Keywords** mHealth · Wellness · Activity monitoring · Diet monitoring · Healthcare

## 1 Introduction and Related Works

An unhealthy lifestyle is one of the main causes of disease burden and therefore of death in the developed world. The modern urban lifestyle is characterized by insufficient physical activity, junk food and high stress levels, all of which affecting people's well-being. In the long term, this kind of lifestyle can lead to health problems

---

M. Forastiere  
Neatec S.p.A., Via Campi Flegrei 34, 80078 Pozzuoli, Naples, Italy  
e-mail: m.forastiere@neatec.it

G. De Pietro · G. Sannino (✉)  
Institute of High Performance Computing and Networking (ICAR-CNR),  
Via Pietro Castellino 111, 80131 Naples, Italy  
e-mail: sannino.g@na.icar.cnr.it; giovanna.sannino@na.icar.cnr.it

G. De Pietro  
e-mail: depietro.g@na.icar.cnr.it

and diseases, such as excessive weight, obesity, strokes, cardiovascular diseases, and diabetes [1]. Indeed, as reported in this same study, among the ten leading causes of death there are elements like tobacco use, inadequate or excessive nutrition, inadequate aerobic exercise, and excessive alcohol consumption. This shows that behavioral risk factors play an important role in understanding changes in overall mortality trends, as also remarked in [2].

The examination of behavioral risk factors, the evidence of their causal relationship with diseases, their prevalence in the population and their combined effects have led researchers in medicine to the conclusion that correcting a few basic behaviors and lifestyle problems could have a major positive effect on the health of the population [3].

Several initiatives have been undertaken aimed at safeguarding and improving long-term health and promoting healthy ageing. Researchers and developers are strongly involved in realizing wellness applications [4–8], ranging from heart-rate activity monitoring, and step counters, to applications for the tracking of eating habits and physical activity.

The assumption made in this work is that people need to consistently follow a certain course of action, including recommendations and suggestions that are suited to their specific case. A “personal trainer”, as they say, is not surprisingly a specialist much in demand today. The ambition of this app is to provide people with something similar, that is, an application that, after collecting a range of information on the individual, is able to “classify” her/him based on her/his individual characteristics (physical parameters and lifestyle) and then to propose specific recommendations to improve her/his well-being. By monitoring the evolution over time of these individual characteristics, the application can also give feedback on the effectiveness of the measures and therefore provide positive stimuli to motivate the user to continue the path taken.

## 2 The Wellness App

The mobile application developed provides several mechanisms for the daily recording of various wellness parameters. The app allows the user to monitor and improve her/his lifestyle by using appropriate indicators that measure the type of nutrition, amount of physical activity, and sleep quality. The application uses an inference reasoning module to provide suggestions that direct the user toward information about the adoption of Mediterranean-type diet, recognized as the best diet for the prevention of cardiovascular disease, an improved physical activity plan and a better quality pattern of sleep.

In detail, the wellness app is able to measure the following indicators:

1. the Mediterranean Adequacy Index (MAI) [9] that measures the type of diet in relation to the Mediterranean diet, considered ideal by many researchers;

2. the Target Calorie Index (TCI) [10] that measures the amount of calories ingested in relation to the optimal value for the individual calculated according to her/his anthropometric data;
3. the Time spent per week on Physical Activity (TPA) [11] that measures the amount of physical activity in comparison with the WHO guidelines; and
4. the Pittsburgh Sleep Quality Index (PSQI) [12] that indicates the quality of sleep.

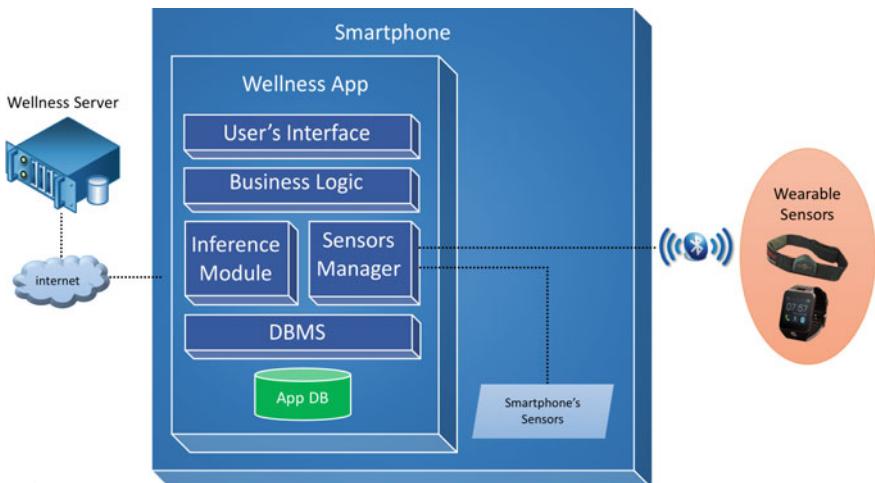
Finally, based on these four indicators, a Wellness Index (IW) is calculated that provides a measurement, albeit partial and incomplete, of the state of physical well-being of the individual.

This latter index is calculated once a week and depends, as shown in Eq. 1, on the mean value of the MAI index values obtained over the week, the mean value of the TPA index values over the same week, the mean value of the TCI index values, and the mean value of the PSQI index values.

$$IW = \frac{mean(MAI) * 2 + mean(TPA) * 2 + mean(TCI) + mean(PSQI)}{6} \quad (1)$$

## 2.1 Software Details

In this section, we will provide an overview of the software architecture designed for our Java-based mobile application shown in Fig. 1.



**Fig. 1** Software architecture of the wellness app

The **User's Interface module** is a software layer external to the business module. This layer contains all the interaction interfaces provided to the user and the mapping phase of these interfaces with the business layer thereby realizes the Model View Controller (MVC) design pattern.

The **Business Logic module** enables the operating of the application and includes the processing core. It provides all the logic required to manage the diet diary, physical activity diary, the user profiling, the privacy logic and the generation of the Wellness Index. It also provides the mechanism that manages the exchange of information between the source app data, through the module DBMS, and the user interface, through the presentation logic, and finally the intermediate processing on the extracted data.

The **Inference module** manages the cycle of reasoning and acts as a main interface between the system of reasoning and the business module. Its mandate is necessary to invoke some internal components, in order to ensure the smooth flow of the execution of inference, the updating of knowledge, and the reporting of the results to the external components of the system. Through this module functions are implemented to provide indications to the user with regard to all monitored parameters and inputs.

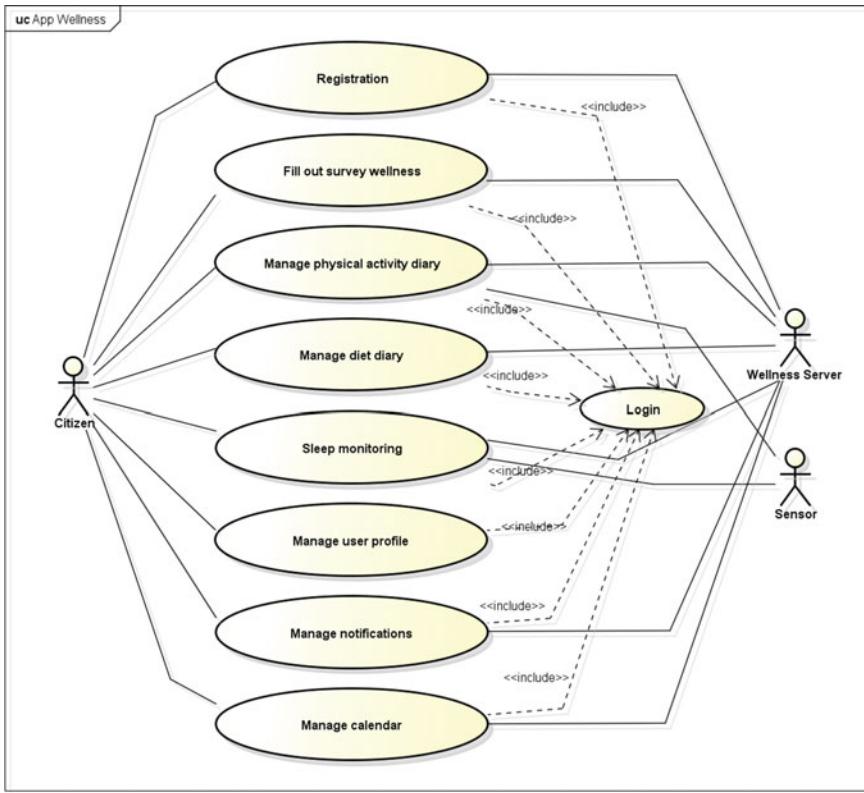
The **Sensor Manager module** deals with the dialogue between the app and the available sensors by creating a virtual layer and allowing the application with different sensors. It is responsible for managing the data exchange via the Bluetooth communication protocol with sensors, such as wristbands, smart watches, or smart elasticised chest-belts, as for example the Zephyr BioHarness.

The **DBMS module** manages the database and the persistence of the Wellness data. Moreover, it provides also mechanisms for the exchange of information between the app and a wellness server, through the use of apposite REST services, for data synchronization, historic data recovery and the updating of the user's goals.

The software has been realized in order to provide the best performance to the user and to be maintainable, dependable and usable. To achieve this, we have followed the software engineering principles, starting from the requirements elicitation, analysis, and specification, summarized in a formal document according to standard guidelines [13]. Subsequently, to design the app, we have used the so-called unified modelling language (UML), a standard notation to visually describe and interpret object-oriented software applications [14]. We have realized UML class diagrams with the use of case descriptions, the entity relationship diagram, the component diagram and the class diagrams. For reasons of brevity, we report in this paper only a use case diagram, Fig. 2, that illustrates the main functionalities provided by the app.

## 2.2 Implementation Details

The application is implemented to be runnable on resource limited devices, like smartphones or tablets. The prototype implementation is entirely based on the powerful Java technology in order to take advantage of its inherent platform-independence.

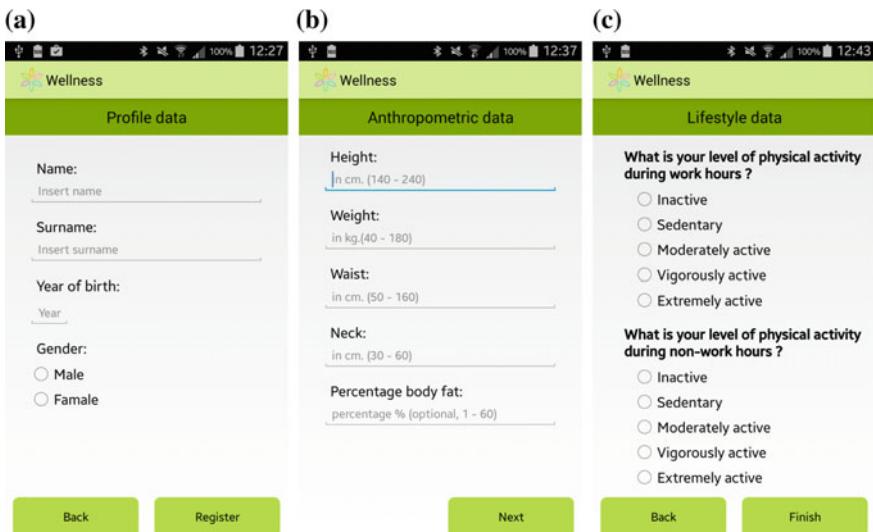


**Fig. 2** Use cases diagram of the wellness app

In terms of the interface design and usage, the wellness app should comply with criteria similar to those of web site development. The design of an aesthetically pleasing interface is important, however, as the success of the system is based on accessing information in an intuitive and easy way [15].

In collaboration with the expert team of the University of Magna Graecia of Catanzaro (Italy), a methodology has been identified to investigate the habits and lifestyle behaviors of a user with reference to the following areas: general information including the anthropometric parameters of the individual, profession, physical activity, nutrition, smoking, and alcohol consumption. The assessment, arising from this information, will be the result of a questionnaire that is submitted to the user. Some screenshots of the questionnaire are shown in Fig. 3.

These data are needed to personalize the recommendations to improve the well-being of the specific user. It is important to note that some of the shown information in the figures are mandatory and must be inserted from the user, as for example the height, the weight, the waistline, and the neck size. Other kind of information, like



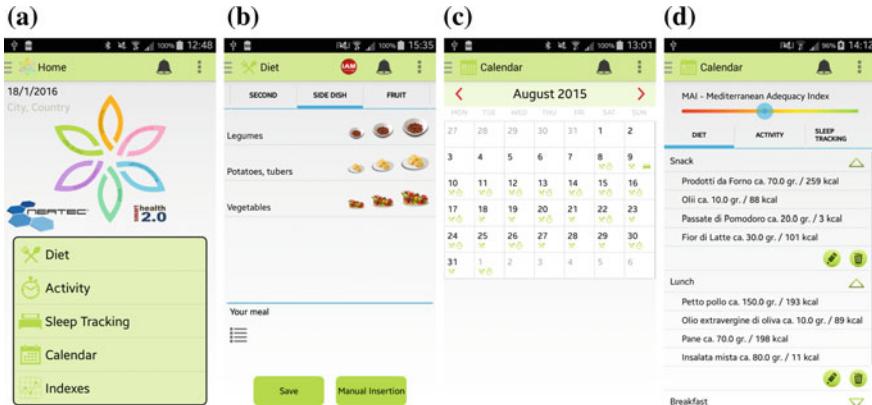
**Fig. 3** User's Data: **a** General information; **b** Anthropometric information; **c** Lifestyle information

the Body Fat Percentage, are automatically estimated by the app starting from the previous mandatory information [16].

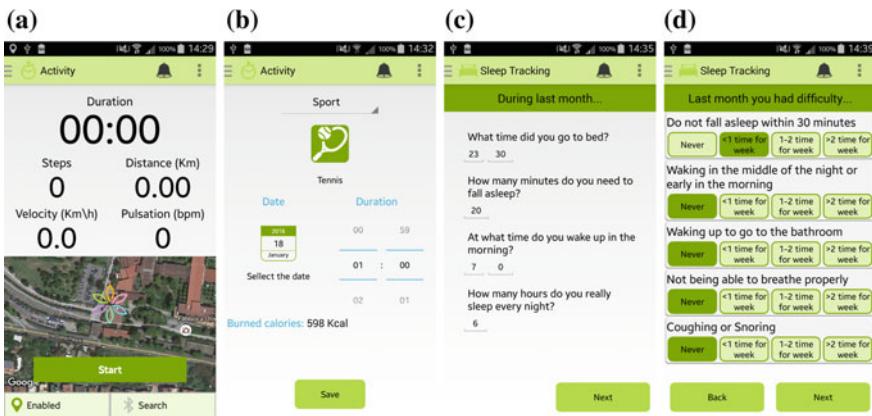
Once the user has entered her/his information, she/he can access the main menu of the wellness app. From the main menu, shown in Fig. 4a, the user can select the option to insert a new meal. By using the simple layout appositely developed, the user can easily inset her/his meal by selecting the type of food and the quantity. Some pre-set quantities have been provided to speed up the insertion, as shown in Fig. 4b. Moreover, some examples of types of food that can be inserted in the app as meals are reported in Table 1. Every day and every week the app automatically calculates the Mediterranean Adequacy Index (MAI) based on the inserted meals. The MAI index assesses how close the diet is to its Reference Mediterranean Dietary Pattern. This index is obtained as illustrated in the study [17].

The user can visualize all inserted data by selecting the Calendar button from the main menu. As shown in Fig. 4c, d, the user can have a monthly view or a daily view where she/he can check, modify or delete the added information.

Furthermore, the user can select the option Physical Activity to insert an activity. She/he can decide to automatically add a movement by using the inertial sensors of the smartphone or an external sensor, such as a smart band, connected to the app by a Bluetooth connection. As illustrated in Fig. 5a, the app is able to monitor the performed steps, the travelled distance (km), the velocity (km/h) and the Heart Rate (bpm) if provided. Moreover, the user can also insert manually the activity, as shown in Fig. 5b by selecting the type of the activity, the date when it is performed and the duration. Some examples of types of activity that can be inserted in the app are reported in Table 1. Finally, the app permits the user to estimate her/his sleep quality



**Fig. 4** Screenshots of the app: **a** Main menu; **b** Insertion of a new meal; **c** Calendar; **d** Details



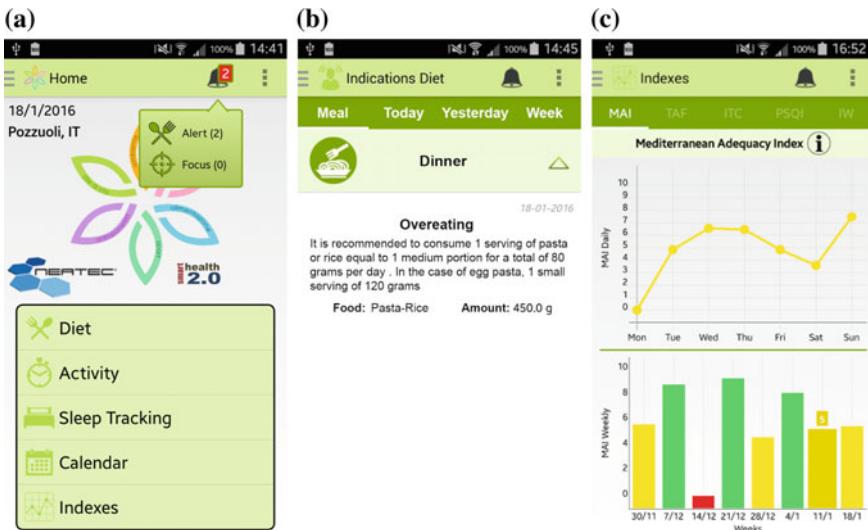
**Fig. 5** Screenshots of Activities and Sleep Monitoring: **a** Automatic insertion of a new activity; **b** Manual insertion of a new activity; **c** and **d** PSQI questionnaire for the sleep quality estimation

by completing the PSQI questionnaire, one of the most popular tools for assessing the quality of sleep [12], as shown in Fig. 5c, d.

Every day, or every time the user inserts any new data, the app automatically calculates the indicators listed in Sect. 2. Every time the app detects an anomaly situation, it promptly alerts the user, as shown in Fig. 6a, b. By clicking on the alert, the user can visualize the recommendations to follow to improve her/his wellness and lifestyle. It is important to note that all the recommendations provided by the app are based on expert knowledge formalized in strict collaboration with the medical team of the University of Magna Graecia of Catanzaro (Italy). Finally, all calculated indexes are available by selecting the Index button from the main menu of the app, as for example, for the MAI index in Fig. 6c.

**Table 1** Examples of information that can be inserted in the wellness app

Category	Typology	Unit
Food	<b>Vegetables and Vegetables:</b> Lettuce, Broccoli, Cucumber, Pumpkin, Zucchini, Pepper, etc.	gr.
	<b>Legumes:</b> Beans, Peas, Chickpeas, Fava beans, etc.	gr.
	<b>Milk:</b> Goat milk, Whole milk, Skimmed milk, etc.	gr.
	<b>Cheeses, dairy product:</b> Fior di Latte, Fontina, Cheese, Gorgonzola, Mascarpone, Smoked provola, Mozzarella, etc.	gr.
	<b>Bread:</b> Wholegrain, White bread, Rye bread, etc.	gr.
	<b>Meat :</b> Chicken, Turkey, Lam, Calf, etc.	gr.
	<b>Fish:</b> Carp, Mullet, Grouper, Cod, etc.	gr.
	<b>Oils and Condiments:</b> Olive oil, Seed oil, Extra virgin olive oil, Butter, etc.	gr.
	<b>Jams:</b> Apricot, Cherries, Raspberry, Apple, etc.	gr.
	<b>Sweets:</b> Chocolate cake, Honey, Chocolate pudding, Meringue, etc.	gr.
	<b>Alcoholic beverages:</b> Wine, Champagne, Beer, Gin, Vodka, etc.	gr.
	<b>Cereals and cereal products:</b> Rolled Oats, Cornflakes, Corn, Rice, Pearl barley, Rye, etc.	gr.
	<b>Fruit juices:</b> Pomegranate, Lemon, Orange, Grapefruit, Apple, etc.	gr.
	<b>Soups:</b> Beef or chicken soup, Vegetable soup, etc.	gr.
	<b>Fresh fruit:</b> Melons, Grapefruit, Prunes, Grapes, Pineapple, etc.	gr.
	<b>Yogurt:</b> Whole milk yogurt, Low fat milk yogurt, Fruit Yogurt, etc.	gr.
Activity	<b>Bycicle:</b> Cycling, Spinning	hh:mm
	<b>Fitness:</b> Push ups, sit-ups, pull-ups, squats, circuit training, body building, yoga, Pilates, stepper, etc.	hh:mm
	<b>Dance:</b> Artistic gymnastics, jazz, hip hop, ballroom dancing, Caribbean dances, etc.	hh:mm
	<b>Daily Activities:</b> Houseworks, gardening, crafts, etc.	hh:mm
	<b>Jogging:</b> Race, track, stroke, etc.	hh:mm
	<b>Sports:</b> Football, golf, horseback riding, rugby, roller-skating, squash, tennis, etc.	hh:mm
	<b>Walking</b>	hh:mm
	<b>Water Activities:</b> Diving activities in general, windsurfing, etc.	hh:mm
	<b>Winter Activities:</b> Ice skating , skiing, cross-country skiing, etc.	hh:mm



**Fig. 6** Screenshots of alerts and of statics of calculated indexes: **a** Main menu of the app where is notified an alert; **b** Details view of the alert; **c** Daily and weekly trends of the Mediterranean Adequacy Index (MAI)

### 3 Conclusions

Behavior and lifestyle are key determinants of health, disease, disability and premature mortality. There is important evidence that demonstrates that unhealthy behaviours increase the risk of the onset of many diseases and therefore could be considered among the causes of the disease itself. In this paper a mobile application for a personalized monitoring of wellness is presented.

The ambition of this app is to provide people with something similar to a personal trainer, an application that, after collecting a range of information on the individual, is able to classify her/him based on her/his individual characteristics (physical parameters and lifestyle) and then to propose specific recommendations to improve her/his well-being. By monitoring the evolution over time of these individual characteristics, the application can give feedback on the effectiveness of the measures and therefore provide positive stimuli to motivate the user to continue the path taken.

A preliminary experimental phase has demonstrated the validity of the presented app, in terms of its stability, usability, and facility of use. Additionally, a trial phase of the app has been carried out in four districts of the Calabria Region (Crotone, Cosenza, San Giovanni in Fiore and Lamezia Terme). We are now involved in the analysis of the data collected during the trial to investigate and demonstrate the real validity of the app.

**Acknowledgments** The authors gratefully acknowledge support from Neatec S.p.A. and from the project Smart Health 2.0 (PON04A2\_C). A special thank to Prof. Agostino Grassi, world renowned nutritionist and secretary of the Mediterranean Diet Foundation. Moreover, the authors wish to thank Prof. Agostino Gnasso, Prof. Francesco Perticone, Dr. Sofia Miceli, and all research group at the University Magna Graecia of Catanzaro (Italy) involved in the SmartHealth 2.0, for their useful contribution to this study.

## References

1. Jenkins, C.D.: Building Better Health: A Handbook of Behavioral Change, vol. 590. Pan American Health Org (2003)
2. Cutler, D.M., Lange, F., Meara, E., Richards-Shubik, S., Ruhm, C.J.: Rising educational gradients in mortality: the role of behavioral risk factors. *J. Health Econ.* **30**(6), 1174–1187 (2011)
3. World Health Organization: 2008–2013 action plan for the global strategy for the prevention and control of noncommunicable diseases: prevent and control cardiovascular diseases, cancers, chronic respiratory diseases and diabetes (2009)
4. Asselin, R., Ortiz, G., Pui, J., Smailagic, A., Kissling, C.: Implementation and evaluation of the personal wellness coach. In: 25th IEEE International Conference on Distributed Computing Systems Workshops, 2005, pp. 529–535. IEEE (2005)
5. Consolvo, S., McDonald, D.W., Toscos, T. et.al.: Activity sensing in the wild: a field trial of ubikit garden. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1797–1806. ACM (2008)
6. Lamminmaki, E., Parkka, J., Hermersdorf, M., Kaasinen, J., Sampasalo, K., Vainio, J., Kolari, J., Kulju, M., Lappalainen, R., Korhonen, I.: Wellness diary for mobile phones. In: Proceedings of the EMBEC '05 (2005)
7. Sannino, G., De Pietro, G.: An evolved ehealth monitoring system for a nuclear medicine department. In: Developments in E-systems Engineering (DeSE), pp. 3–6. IEEE (2011)
8. Coronato, A., De Pietro, G., Sannino, G.: Middleware services for pervasive monitoring elderly and ill people in smart environments. In: 2010 Seventh International Conference on Information Technology, pp. 810–815. IEEE (2010)
9. Fidanza, A., Fidanza, F.: Mediterranean adequacy index of Italian diets. *Public Health Nutr.* **7**(07), 937–941 (2004)
10. Rising, R., Harper, I.T., Fontvieille, A.M., Ferraro, R.T., Spraul, M., Ravussin, E.: Determinants of total daily energy expenditure: variability in physical activity. *Am. J. Clin. Nutr.* **59**(4), 800–804 (1994)
11. Lee, S.M., Burgeson, C.R., Fulton, J.E., Spain, C.G.: Physical education and physical activity: results from the school health policies and programs study 2006. *J. Sch. Health* **77**(8), 435–463 (2007)
12. Buysse, D.J., Reynolds, C.F., Monk, T.H., Berman, S.R., Kupfer, D.J.: The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* **28**(2), 193–213 (1989)
13. Robertson, J., Robertson, S.: Volere. Requirements Specification Templates (2000)
14. Ambler, S.W.: The Unified Modeling Language v1. 1 and Beyond: The Techniques of Object-Oriented Modeling. An AmbySoft Inc. White Paper. <http://www.ambysoft.com/umlAndBeyond.html>
15. Tractinsky, N., Katz, A.S., Ikar, D.: What is beautiful is usable. *Interact. Comput.* **13**(2), 127–145 (2000)
16. Siri, W.E.: The gross composition of the body. *Adv. Biol. Med. Phys.* **4**(239–279), 513 (1956)
17. Alberti-Fidanza, A., Fidanza, F., Chiuchiù, M. P., Verducci, G., Fruttini, D.: Dietary studies on two rural Italian population groups of the Seven Countries Study. 3. Trend of food and nutrient intake from: to 1991. *Eur. J. Clin. Nutr.* **53**(11), 854–860 (1990)

# “White Coat” Effect Study as a Subclinical Target Organ Damage by Means of a Web Platform

J. Novo, A. Hermida, M. Ortega, N. Barreira, M.G. Penedo,  
J.E. López and C. Calvo

**Abstract** “White-coat” effect designs those hypertensive subjects with “uncontrolled” office Blood Pressure (BP) but normal BP values when assessed by Ambulatory BP Monitoring (ABPM) or home BP monitoring (HBPM). Cardiovascular Risk (CV) risk is lower than those with real uncontrolled BP but it still remains unclear if it is comparable to those well controlled hypertensive subjects. This paper presents the study, the results and the web platform that was designed and implemented which make possible the study of the “White-coat” effect as a subclinical target organ damage. The large amount of information that needs to be gathered, calculated and analyzed makes specially complicated, even almost inviable, the development of the study by traditional manual routine. This motivated the implementation of a platform that permitted the doctors make the study with guarantees. The implemented web platform organizes the information of the patients, presenting all the necessary information for the study, including physical and clinical parameters or 48-h ABPM processing. It also automatically estimates glomerular filtration rate by MDRD equation (GFR) and Sokolow-Lyon criteria for left ventricular hypertrophy (LVH) as well

---

J. Novo (✉) · M. Ortega · N. Barreira · M.G. Penedo  
Departamento de Computación, Universidade da Coruña, A Coruña, Spain  
e-mail: jnovo@udc.es

M. Ortega  
e-mail: mortega@udc.es

N. Barreira  
e-mail: nbarreira@udc.es

M.G. Penedo  
e-mail: mgpenedo@udc.es

A. Hermida · J.E. López · C. Calvo  
Complexo Hospitalario Universitario de Santiago de Compostela,  
Universidade de Santiago de Compostela, Santiago de Compostela, Spain  
e-mail: Alvaro.Hermida.Ameijeiras@sergas.es

J.E. López  
e-mail: joseenriquelopezpaz@yahoo.es

C. Calvo  
e-mail: igamevas.calvo@usc.es

as different statistics from the 48-h ABPM. The platform facilitates the doctor's work avoiding large and tedious manual processes, minimizing the risk of possible miscalculations and analyzing all the information in a easier way. This framework helped the doctors to recognize the so called "ABPM effect", and what is more important in the management of hypertensive subjects, it helps to better identify hypertensive subjects at poor cardiovascular prognosis.

**Keywords** Medical informatics applications • Web-based systems • Internal medicine • Hypertension • Ambulatory blood pressure monitoring • White coat effect

## 1 Introduction and Previous Work

Cardiovascular disease (CVD) is the main cause of death among developed societies over the world as the statistics demonstrates, for example, in the US [1], stating that CVD is the main cause of death among adult population. Moreover, CVD is the first cause of losing Disability Adjusted Life Years (DALYs), being the 23 % of all the cases, what represents a higher number than neoplasias (17.9 %) and the neuropsychiatric conditions (16.5 %) [2]. It also represents one of the main causes of years of life lost (YOLL). In the particular case of Spain, it was the third cause of YOLL by preventable deceases more important (16 % of all) after neoplasias and caused more than 5 million of hospitalizations in 1996, for instance, that is, more than the 12 % of the total registered that year [3].

BP monitoring is a crucial task as it measures its state and a hypothetical degree of hypertension and, therefore, the associated risk factor with a possible CVD. Traditionally this step was performed with the capture of single BP measurements in specific moments as, for example, in the clinical revision by the doctor. This process has two main drawbacks: first, the large variability in the measurements depending on the moment of the clinical revision and the situation of the patient (moment of the day, stress, etc.) and second, the influence of the doctor due to its presence (for example, inducing nervousness), called "white coat" (WC) effect, that varies the results, hiding the real state of the patient.

For that reason, the ambulatory BP monitoring (ABPM) can be a more reliable and complete measurement set to analyze the BP condition of the patient. The ABPM is an outpatient monitoring that helps to track the BP evolution of the patient in a period of time. This monitoring process is performed with the holter monitor, that is a device for the continuous monitoring of the blood pressure. With the use of the ABPM doctors obtain more accurate information, being able to analyze the evolution of the BP in large periods including day and night measurements and, therefore, guiding to better analysis and diagnosis of the patients. ABPM is a more sensitive risk predictor of clinical CV outcomes than office BP as it has been evidenced from meta-analyses of published observational studies [4]. It is currently considered the reference for out-of-office BP and the most accurate method for confirming a diagnosis of hypertension.

The Center for Medicare and Medicaid Services in the United States approved ABPM for reimbursement for the identification of subjects with white-coat hypertension [5] and the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom recommended that ABPM should be offered as a cost-effective technique to all people suspected of having hypertension [6].

WC effect indicates those hypertensive subjects with “uncontrolled” office BP but normal BP values when they are measured by ABPM or home BP monitoring (HBPM), whereas white-coat hypertension exists if the office BP is high, and the awake ambulatory BP is normal in a patient not receiving antihypertensive medication [7, 8].

The overall prevalence of white-coat hypertension averaged 13 % (range 9–16 %) and it amounted to about 32 % (range 25–46 %) among hypertensive subjects. It still remains unclear whether subjects with WC hypertension can be equaled to true normotensive from a CV prognosis point of view [9]. Although target organ damage and CV event rates are less prevalent in WC hypertension than in sustained hypertension, patients with WC hypertension are in a higher chance of developing sustained hypertension and they should be scheduled for regular follow-up to check the CV risk profile and searching for asymptomatic target organ damage [10].

During the last years, with the spread of the technology, the computer applications increased in relevance over the majority of the health-care systems. Thus, they have become an important utility to provide and support health-care procedures where, for example, distances are a drawback [11]. In this case, they help to gain efficiency in medical treatments and facilitating the collaboration among different specialists, reducing costs and the necessary resources [12]. Additionally, the automatization of medical procedures can help the specialists to speed up the patient diagnosis, facilitating the processing of large databases that any health-care system needs to handle. This is one of the main properties that any medical application should present aiming to provide real time, replacing manual procedures that are tedious and highly time consuming. Moreover, automatic medical applications imply objective and repeatable methodologies, overcoming the subjectivity introduced by any specialist and, therefore, increasing the quality of the diagnosis and treatment of the patients.

Medical computer applications can also help the doctors to gather large and heterogeneous patients data for clinical trials or medical studies, studies that otherwise could be highly complicated or impossible to reproduce. In that sense, a web application was constructed that permitted the doctors include all the necessary information and study the influence of the “white coat” in possible hypertensive patients analyzed with the ambulatory BP monitoring. To date, we are not aware of any application which let specialists assess differences on incidence on target organ damage between hypertensive subjects regarding ABPM registers, as the web platform that was implemented permitted.

## 2 The Study

A retrospective cohort study was designed in order to assess the prognostic CVR value of WC effect regarding subclinical target organ damage (TOD) in a cohort of hypertensive subjects.

In order to include enough information for the study, a cohort with 344 hypertensive patients was analyzed and their results were included in the study. They all underwent a complete annual medical examination from January 2011 to December 2014, including 48-h blood pressure monitoring (ABPM) with a Spacelabs 90207 device adjusted automatically monitoring every 20 min between 07:00 am and 23:00 pm and every 30 min at night.

Given the large amount of information that should be analyzed as the study include a large set of heterogeneous variables, the manual procedure would be largely complicated and tedious even being, therefore, the study viability at risk. Thus, the possibility of having a tool that facilitate the gathering of all the necessary data, automatically calculating all the derived parameters and presenting all the information to the specialists would be ideal.

This situation motivated the development of a platform that facilitates the doctor's work, achieving the objectives of the study with guarantees. The implemented website platform was used in a two-step process. Firstly, by analyzing ABPM along 48 h and identifying those subjects with WC effect. Secondly, the platform was used to record and analyze results from the following examinations: intima/media thickness (IMT), carotid femoral pulse wave velocity (PWV) (by Sphygmocor At Cor<sup>®</sup>) and oscillometric measurement of ankle-brachial index (ABI). It also allowed us to estimate automatically glomerular filtration rate (GFR) by MDRD equation and Sokolow-Lyon criteria for left ventricular hypertrophy (LVH).

## 3 The Web Tool

A research framework was designed and implemented to aid the doctors to make a complete analysis of the patients of a cohort and determine the hypothetical WC effect as a subclinical TOD of hypertensive patients. The framework development was performed under the supervision of the specialists in order to guarantee that the website platform include all the information and characteristics that are necessary to do the study.

The design and implementation was performed in collaboration with doctors in an interactive process, being them who tested and sent feedback in order to correct and complete the tool. The collaboration helped to identify the main characteristics that the platform should present for the study. The entire platform, both architecture and contents, was designed having in mind the following required premises:

- The platform should be accessible at any time to facilitate the study given the dispersion of the doctors and their access to the data from different workstations.
- Under the clinical point of view, the framework should provide a complete medical profile for each patient, organizing the clinical information in progressive revisions that facilitate the patient analysis and evolution over the time.
- Given the large amount of data associated to any patient, this information should be presented graphically in the most intuitive and easiest possible way in order to facilitate the doctors work and minimizing the risk of mistakes.

A web application was implemented, instead of a standalone app, that is running in a central server where is stored all the data. Hence, any doctor, being physically in different clinical units, can access directly to the platform from the corresponding workstation and work with the platform without needing a specific installation. It is only necessary a web browser and Internet connection to access to the platform as any other website. The web application can also be updated over the time, refining functionalities or including new ones. In this case, the web deployment can be updated in the central server and the specialists would have access to the new version directly. As the platform is centralized, all the data is centralized as well, letting further statistic analysis of the data could be made.

To guarantee the independence between the user’s interaction and the implementation, a model-view-controller (MVC) design pattern was the most suitable option. MVC is represented by 3 different layers, each one with an specific purpose:

- **Model layer:** responsible for the storage and management of all the data that is included in the application.
- **View layer:** generates the output representation, that is, the way the users interacts with the application. The main hypertension services are implemented in this layer, keeping in the database only the original patient data.
- **Controller layer:** this layer is in charge of the communication between model and view layers, converting data formats to be handled from model to view, and vice versa.

### **3.1 Register Patient and Checkup Modules**

The first, and main, functionality that is included in the platform consists of the registration of new patients in the system and/or adding a new patient revision. Some parameters are stable for a patient, being the rest linked to a particular revision. Hence, the patient registration firstly introduce the general unalterable information: identification code, birthday (to derive the age), gender or ethnicity. Last three are necessary in the study. Afterwards, the module moves forward to introduce the corresponding first checkup, which is the same as any further checkup. Figure 1 illustrates the register patient page.

Hola Medico

Insertar paciente - Inicio - Salir

**Usuarios**

Cambiar contraseña

**Médicos**

Insertar paciente  
Buscar paciente  
Listar pacientes

Comprobar fichero

**Insertar paciente (1/3)**

**Datos generales**

Código de paciente: [Input field]

Fecha de nacimiento: 01 01 1920

Sexo:  Hombre  Mujer

Raza: Caucásica

Antecedentes familiares de enfermedad cardiovascular prematura

Siguiente

**Fig. 1** Register patient module, birthday, gender and Ethnicity variables, to be filled for the study

The checkup module, as indicated by the specialists, contains the list of variables and parameters that are needed in the posterior analysis of the study. The module include the following information:

- **Physical condition:** includes weight, waist circumference, height, and derives the corresponding body mass index.
- **Manual blood pressure measurement:** taking the systolic and diastolic BP with both arms, seated and standing for a more reliable measurement.
- **Electrocardiogram:** including the different parameters of the electrocardiogram.
- **Other measurements:** in addition, other necessary parameters for the study are also introduced: the IMT, ABI and PWV.

Figure 2 show the physical and clinical conditions form with some of the parameters that have to be filled. With all this information the platform automatically calculates the glomerular Filtration rate and the left ventricular hypertrophy. It also identifies the patients with target organ damage, which facilitates the analysis and diagnosis of hypertensive patients with WC effect.

**ABPM Analysis Module** The checkup module also includes another functionality, the ABPM analysis to track the BP in periods of 48 h. This performs a more complete and reliable set of information regarding BP than single measurements in the clinic.

Here, the specialists can upload the records that were acquired with the ABPM, and being processed by the platform. This information is presented graphically to the user, for a better inspection. This graphic includes blood pressure systolic, diastolic and pulse. An example of a graphic representing the evolution of the blood pressure of a patient over 48 h is presented in the Fig. 3.

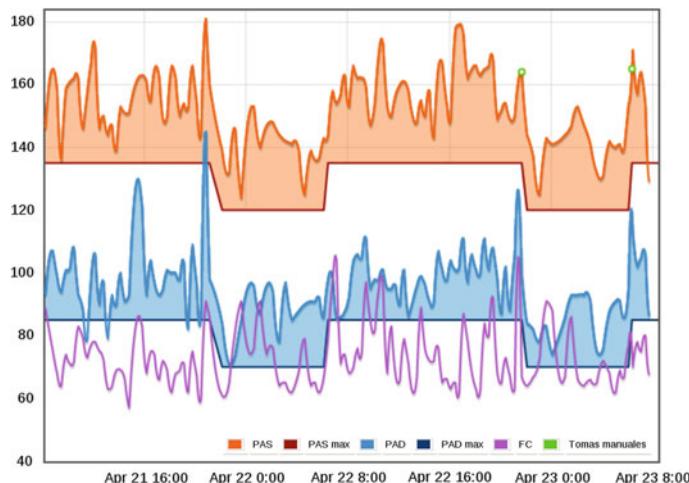
The platform automatically calculates the following parameters that are used in the WC effect study: 48-h systolic BP, 48-h diastolic BP, nocturnal systolic BP falling and nocturnal diastolic BP falling.

**Datos antropométricos**

Peso	<input type="text"/>	kg
Talla	<input type="text"/>	cm
IMC	<input type="text"/>	
Perímetro de cintura	<input type="text"/>	cm
Índice Talla/cintura	<input type="text"/>	
Perímetro del brazo dominante	<input type="text"/>	cm
Utilización de manguito de obeso	<input type="checkbox"/>	

**Medidas presión arterial**

Hora de la medida	11	29	
Tratamiento antihipertensivo previo	No		
<b>Brazo izquierdo</b>			
	PAS	PAD	FC
Sentado 1	<input type="text"/> 153.0	<input type="text"/> 76.0	<input type="text"/> 58.0
Sentado 2	<input type="text"/> 138.0	<input type="text"/> 71.0	<input type="text"/> 57.0
Sentado 3	<input type="text"/> 133.0	<input type="text"/> 71.0	<input type="text"/> 58.0
Media	<input type="text"/> 148.0	<input type="text"/> 77.0	<input type="text"/> 61.0
De pie	<input type="text"/>	<input type="text"/>	<input type="text"/>

**Fig. 2** Physical and BP manual measurements. Variables, that were used in the study, to be filled**Fig. 3** Blood pressure and pulse evolution of a patient during a period of 48 h, example of platform representation. x-axis, time of the day. y-axis, blood pressure. PAS: Systolic Blood Pressure. PAS max: Maximum Systolic Blood Pressure considered as normal. PAD: Diastolic Blood Pressure. PAD max: Maximum Diastolic Blood Pressure considered as normal

## 4 Results

The framework was implemented using different technologies that permitted the inclusion of all the mentioned characteristics. In that sense, the application was implemented using Java 2 Enterprise Edition architecture using the Struts library to implement the Model-View-Controller design pattern. Web pages were created with XHTML, CSS and JavaScript. Apache Tomcat was used as the web server and PostgreSQL as the relational database.

The validation of the framework was done in a real scenario, under the use of the Hypertension Unit of the Health Complex of Santiago de Compostela, unit that evaluated its well functioning in a previous testing stage before performing the data gathering for the study, demonstrating its suitability in clinical practice and for the study. To date, there is no other application that include all the characteristics that are needed to succeed with a study of these characteristics.

The web platform was used to record and analyze 48-h ABPM in order to diagnose those with WC effect. As stated, thanks to the platform, doctors could analyzed the results obtained for the two consecutive days and identify the mentioned “white coat” in the first period of 24 h. 344 hypertensive patients were studied from January 2011 to December 2014, a number to be able to obtain conclusions with guarantees.

Furthermore, results from following examinations were also recorded and analyzed: carotid artery ultrasound with intima/media thickness, carotid femoral pulse wave velocity by Sphygmocor At Cor® and oscillometric measurement of ankle-brachial index. This website application allowed us to estimate automatically the glomerular filtration rate by MDRD equation and Sokolow-Lyon criteria for left ventricular hypertrophy.

In the study, by using this platform, we detected differences between those hypertensive subjects with WC effect and those “well-controlled” (21.7 vs. 7.7 % on LVH incidence rate and 11.3 vs. 9.5 m/s on PWV). On the other hand, as this application let us register several “follow-up” visits; it was possible to evaluate estimated risk of any target organ damage along the study period (over 60 % when WC effect was present vs. 41.1 % on well controlled hypertensive subjects).

## 5 Conclusions

This work presents the study of the “White coat” effect as a subclinical target organ damage. This study involved the analysis of a large and complex set of variables, deriving different measurements and implying a tedious and exhaustive process. This procedure that would be otherwise largely complicated, motivated the development of a web platform that helped the doctors to analyze into a quick and easy manner several clinical conditions that lead to high cardiovascular risk,

This web platform also avoids possible miscalculations, thanks to the automatic measurement of several risk functions and medical formulae (i.e. MDRD equation

to glomerular filtration rate estimation or Sokolow-Lyon criteria for left ventricular hypertrophy) along with the automatic assessment of those subjects with apparently “uncontrolled” office BP but normal BP values when assessed by ABPM (“white-coat effect”). To date, there is no other application which let the specialists assess the differences on incidence on target organ damage between hypertensive subjects regarding ABPM registers.

We can conclude that those hypertensive subjects with WC effect have higher incidence rate of subclinical TOD than those with well controlled blood pressure.

**Acknowledgments** This work is supported by the Instituto de Salud Carlos III of the Spanish Government and FEDER funds of the European Union through the PI14/02161 and the DTS15/00153 research project.

## References

1. Go, A.S., Mozaffarian, D., Roger, V.L., et al.: Heart disease and stroke statistics-2014 update: a report from the American Heart Association. *Circulation* **129** (2014)
2. Murray, C.J., Lopez, A.D.: Global mortality, disability, and the contribution of risk factors: global burden of disease study. *Lancet* **349**(9063), 1436–1442 (1997)
3. Defunciones según la causa de muerte, 1999. Instituto Nacional de Estadística. Spain (2002). <http://www.ine.es>
4. Conen, D., Bamberg, F.: Noninvasive 24-h ambulatory blood pressure and cardiovascular disease: a systematic review and meta-analysis. *J. Hypertens.* **26**, 1290–1299 (2008)
5. CMS.gov. Centers for Medicare & Medicaid Services: Medicare coverage policy decisions. abpm monitoring (# cag-00067n), 2001. *J. Hypertens.* (2013)
6. National Institute for Health and Clinical Excellence (NICE): Hypertension: the clinical management of primary hypertension in adults. In: Clinical Guideline, p. 127 (2011)
7. Fagard, R.H., Cornelissen, V.A.: Incidence of cardiovascular events in white-coat, masked and sustained hypertension vs. true normotension: a meta-analysis. *J. Hypertens.* **25**, 2193–2198 (2007)
8. Franklin, S., Thijss, L., Hansen, T.W., O’Brien, E., Staessen, J.A.: White-coat hypertension: new insights from recent studies. *Hypertension* **62**, 982–987 (2013)
9. Pierdomenico, S.D., Cuccurullo, F.: Prognostic value of white-coat and masked hypertension diagnosed by ambulatory monitoring in initially untreated subjects. *Am. J. Hypertens.* **24**, 52–58 (2011)
10. European Society of Hypertension: (ESH) and of the European Society of Cardiology (ESC). 2013 esh/esc guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension. *J. Hypertens.* **31**(7), 1281–1357 (2013)
11. Field, M.J. (ed.): *Telemedicine: A Guide to Assessing Telecommunications in Health Care*. National Academy Press, Washington, DC (1996)
12. Charles, B.: Telemedicine can lower costs and improve access. *Healthc. Finan. Manag.* **54**(4), 66–69 (2000)

**Part VI**

**Smart Medical and Healthcare Systems**

**2016 Workshop**

# **GENESIS—Cloud-Based System for Next Generation Sequencing Analysis: A Proof of Concept**

**Maider Alberich, Arkaitz Artetxe, Eduardo Santamaría-Navarro,  
Alfons Nonell-Canals and Grégory Maclair**

**Abstract** With the advent of the technology, the DNA sequencing has become cheaper and faster. Next-Generation Sequencing platforms are providing new opportunities to address biological and medical issues. However, they present new challenges of storing, handling and processing, as they produce massive amounts of data. Powerful computational infrastructure, new bioinformatics softwares and skilled people in programming are required to work with the analysis tools. This project aims to design and develop an intelligent system that analyses high-throughput datasets, with the purpose of improving the effectiveness in the biological and medical research fields. The target is to make a user-friendly tool that allows the user to automatically or manually design the desired analysis workflow. Therefore, the technological challenges consist in: (i) an interface between clinician and bioinformatics language, (ii) an intelligent tool that selects the appropriate analysis workflow and (iii) a solution that can handle, store and manage big datasets at a reasonable-price. In order to tackle these bottlenecks, a cloud-based prototype enhanced by a graphical user-friendly interface and implemented using Amazon Web Service.

**Keywords** Next generation sequencing (NGS) · High-throughput sequencing · Automatized workflow · Cloud-computing · Amazon web services (AWS)

---

M. Alberich (✉) · A. Artetxe · G. Maclair  
Vicomtech-IK4, San Sebastian, Spain  
e-mail: malberich@vicomtech.org

M. Alberich · A. Artetxe · G. Maclair  
Biodonostia Health Research Institute, San Sebastian, Spain

E. Santamaría-Navarro · A. Nonell-Canals  
Mind the Byte, Barcelona, Spain

## 1 Introduction

Next Generation Sequencing technologies (NGS) or Second Generation sequencers allow sequencing in parallel millions of DNA fragments at unprecedented cost and time, enabling new scientific achievements [1]. There are different types of NGS sequencers, but all follow the same procedure: (i) sample preparation, (ii) sequencing, (iii) analysis of the output, and (iv) validation of the results. This last step is very important as NGS technologies have low accuracy [2, 3]. Second Generation technologies support a wide range of applications, such as whole genome sequencing or gene expression profiling.

Despite of their improvements, the huge amount of data generated transform the analysis step into big data science problems, needing powerful computational infrastructures. As a lot of investment is required, such infrastructures are difficult to deploy in research centers, hospitals and companies. Cloud computing solutions are gathering strength and seem to be a solution to deal with those issues of big data in NGS analysis field at a moderate price [4–6].

As NGS data analysis field is evolving rapidly, plenty of analysis tools have been developed. Users could have difficulties in selecting the most suitable tool for analyzing their data. Moreover, most solutions deal with a single step of the whole analysis. The challenging part is that the users need to select the appropriate set of tools for their data without having a well-documented review of the NGS tools. In addition, most tools require advance programming skills due to the fact that few solutions supply with user interface.

In order to solve the exposed problems, the current paper presents the description and a first implementation of a cloud-based system for NGS analysis workflow design. The aim of this project is to develop a user-friendly system that allows researchers to design automatically or manually the entire analysis pipelines based on existing tools. This work will imply the selection and integration of some of the most used tools in a unified system, so that the user can easily and rapidly choose or design a suitable workflow to analyze the data. The scientific challenges of this work are: (i) the design of Software as a Service (SaaS) platform that handles, stores and analyzes NGS data in a secure and private manner, (ii) development of an intuitive graphical user-friendly interface for all type of users, (iii) integration of different analysis libraries and tools in a transparent way for the user and (iv) automatization of the design of the pipeline for the different analysis workflows. Thereby, the proposed solution will allow researchers to optimize their time and focus in concluding the results of the analysis.

This paper is structured as follows: in Sect. 2, different tools and technologies related to NGS are introduced and compared with the proposed approach. In Sect. 3 the architecture of the system is exposed. Section 4 details the implementation of the first prototype. Finally, in Sect. 5 conclusions and future work are discussed.

## 2 State of the Art

High-throughput data analysis is a complex task, which includes multiple steps. Depending on the objectives of the user, those steps are different. Many current solutions have been designed to deal with a single step analysis. Furthermore, most of the approaches are not user-friendly, since basic or more advanced informatics skills are needed to configure and execute the different analysis steps via command lines or via scripting.

To address these challenges, new tools based on analysis workflow design rise importance. There are several solutions which the principal objective is that the user can manually build, modify, interconnect and execute the desired analysis pipeline. In this section, a comparison of the proposed approach in this paper with the most known workflow systems (Galaxy [7], LONI [8] and Taverna [9]) is done.

Table 1 shows that contrary to the most popular platform, Genesis has been implemented to support all type of second generation sequencers. The four are available for all the Operating System (OS), but only the proposed solution and Galaxy provide cloud-based approach for the analysis. Even if Table 1 points out that all applications have graphical interface (GUI), only some of them are user-friendly supplying with ergonomic and easy-to-use tools for creating analysis pipelines. Furthermore, Genesis is the only system which pretends to automatize pipeline building step with conditions for selecting parameters and algorithms for each type of sequencer. Last difference is that the proposed prototype does not require user's expertise in NGS's bioinformatics data analysis tools and Information Technology (IT) [10, 11], since the system will proposed by default a valid pipeline that depends on the sequencer and what the scientists is looking for.

The current work extends the actual solutions, by adding visual and interactive user interface to easily design analysis workflow, and by adding automatize pipeline design feature.

## 3 Proposed Approach

To fulfill the needs of NGS data analysis, the presented paper proposed to design a SaaS solution using Amazon Web Services (AWS) that offer: (i) easy deployment, (ii) security and access management services, (iii) flexibility and scalability of the

**Table 1** Galaxy, LONI, Taverna and Genesis solutions characteristics [10]

Software	Platforms	OS	GUI	Cloud	Bioinformatics skills	Automatized pipelines
Galaxy	Illumina, SOLiD	All	Yes	Yes	Required	No
LONI	Illumina, SOLiD	All	Yes	No	Required	No
Taverna	Illumina, SOLiD	All	Yes	No	Required	No
Genesis	All	All	Yes	Yes	Not required	Yes

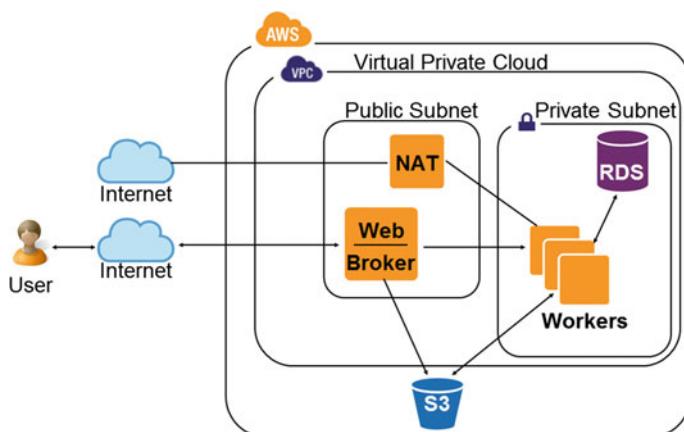
resource (scale on demand), (iv) cost-effective solution (pay-as-you-use), and (v) intuitive and easy-to-use web interface, (vi) automatic analysis workflow design, and (vii) integration of most used bioinformatics libraries in a transparent way for the user.

Figure 1 shows the architecture of Genesis solution. This AWS configuration includes Virtual Private Cloud (VPC) with public and private subnets, increasing security by controlling the traffic at different levels of the architecture. The components to launch and establish in the public subnet are: (i) Elastic Compute Cloud (EC2) with the role of deploying the web application and with the function of broker, which controls the traffic of the different tasks requested by the user, and (ii) a Network Address Translation (NAT) instance to enable traffic from the private subnet to Internet (so that elements of the private subnet can access and download packages from the internet).

The services to configure in the private subnet are: (i) EC2 instances known as workers with the function of performing analysis tasks, and (ii) Relational Database Service (RDS) instance which is the database of the project where information of the analysis stored.

Moreover, Simple Storage Service (S3) is used to store and retrieve files and graphs of the analysis, as it is an object based storage web service offered by AWS.

The prototype work as follow: (i) user logs in the corresponding web address to access our solution, (ii) user selects or designs workflow depending on the working mode desired (automatic or manual), (iii) user uploads any NGS data file, (iv) this file is uploaded to the web instance and simultaneously to S3, (iv) the file is deleted from the web instance to avoid memory problems, (v) when the user request to process the analysis, the different tasks to realize are send to broker instance, (vi) broker is responsible to check availability of the workers, and if it is necessary or not to create a worker instance, (vii) the worker carries out the tasks, and (viii) when any task is finished the corresponding fields of the database are updated and results are placed in S3.



**Fig. 1** Different AWS elements used to design the prototype architecture

## 4 Implemented Solution

Genesis encapsulates all the necessary scripts, packages and tools for analyzing NGS pipelines for different type of sequencers using AWS resources. The project is available online. The users can access Genesis with their credentials from any web browser (Microsoft Internet Explorer, Mozilla Firefox, Google Chrome and Apple Safari), avoiding complications with the installation and updates.

The project is principally implemented in Python programming language. All functionalities offered by the prototype run on Linux servers. As depicted in Fig. 1, there are different components in Genesis architecture. Each element has its own function. The following section gives an overview of the purpose and implementation of the different elements of the prototype.

### 1. Broker Instance

Its main function is to create, translate and manage the analysis tasks requested by the user. It also supervises workers: (i) scale up the necessary number of workers, and (ii) distribute the tasks among the workers. Celery [12] (library that has functions for message parsing between sender and receiver, in this case, worker and broker respectively) and RabbitMQ [13] (responsible to translate messages between sender and receiver) packages are utilized to perform the role of this type of instance.

### 2. Worker Instance

This component of the schema is responsible of performing the different steps of the NGS data analysis. Each step of the analysis process is defined as Celery tasks. This enables developers to add easily new functionalities to the system. Most of the tasks are scripts that wrap most successful NGS data analysis tools. Apart from installing those tools and Celery package, SQLAlchemy [14] (a Python database toolkit) is required since the worker needs to connect with the user database to update the results of analysis process.

As the number of workers varies depending on the requests of the users, an Amazon Machine Image (AMI) service has been used in order to replicate instances with the essential packages and tools. Moreover, the source code related to the tasks is available in S3 repository for every new worker that is created.

Actually, we have implemented the different tasks of analyzing one of the most used NGS data workflow: Re-sequencing pipeline containing Input, Quality Control, Trimming, Alignment, SNP identification and Visualization steps.

Regarding the tasks, Input task is responsible of uploading the file to S3 using Boto package [15] (library with interfaces to AWS). Two different algorithms have been implemented to manage different file size. Additionally, this module extracts parameters (i.e. extension) and in case it is necessary, decompresses the input file.

Quality Control is responsible for assessing the quality of the sequences at different points in the pipeline. Six different metrics are implemented using Fastq package [16] (contains functions for assessing the quality control of the sequences).

Trimming step purpose is to remove more accurate results. Genesis platform proposes eleven different trimming options. The user can select one or more trimming metrics via an HTML multiple choice form (cf. Fig. 2).

For alignment and SNP detection tasks, Genesis platform uses existing tools (Bowtie [17], Samtools [18], VCFTools [19]) using subprocess package [20] (module interface to create and work with additional processes).

Several viewers are available to: (i) visualize the results of the quality control process which uses Matplotlib package [21] (a plotting Python package), and (ii) alignment and SNP detection viewers have been deployed with JBrower [22] (help building genome browser in Javascript and HTML5).

The system proposes an automatic pipeline creation tool depending on the input data. For example, the trimming metrics performs differently depending on the

**Trimming Metrics**

---

**Cut n bases from 3' end of reads**

**Cut n bases from 5' end of reads**  
  
 Remove any N's from the 3' end  
 Remove any N's from the 5' end

**Remove poly A from 3' end**

**Remove poly T from 5' end**  
  
 Cut length to the minimum size

**Remove adaptors from 3' end**

**Remove adaptors from 5' end**

**Remove primer from 3' end**

**Remove primer from 5' end**

---

**Fig. 2** Different trimming options available for the user for analysis configuration

sequencer type: for Roche 454 sequencers, it cuts 20 basepair (bp) from the end of the sequences, and for the Illumina sequencers, it removes all the sequences to the minimum length of the sequences.

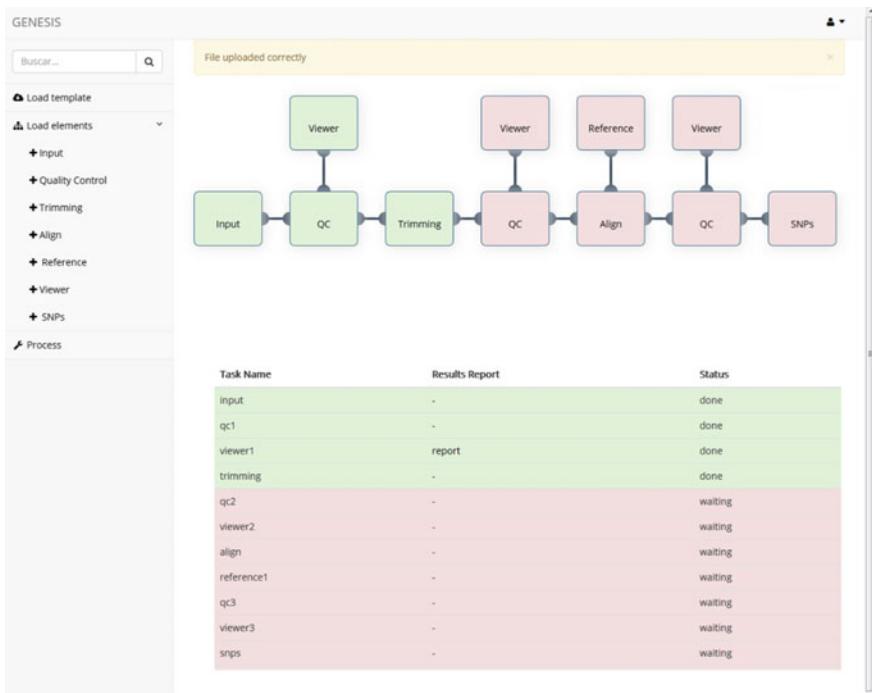
### 3. Web Instance

The deployment of Web instance counts with two steps: (i) design and development of a web-page interface according to Genomic group of BioDonostia Health Research Institute recommendations, and (ii) deploy framework and web server.

#### (a) Web-page Interface

One goal of the prototype is to provide with an ergonomic and intuitive graphical user interface. The final version uses HTML, CSS and Javascript programming languages to obtain a dynamic and interactive web page. Furthermore, different free toolkits have been used for a faster and easier design: (i) Bootstrap [23] templates as base of the web-page, and (ii) jsPlumb [24] toolkit which provides with functions for implementing flowchart and graphical pipeline applications.

Three sections can be distinguished in the aspect of Genesis (Fig. 3): (i) configuration menu with different functions related to the user (top right side menu),



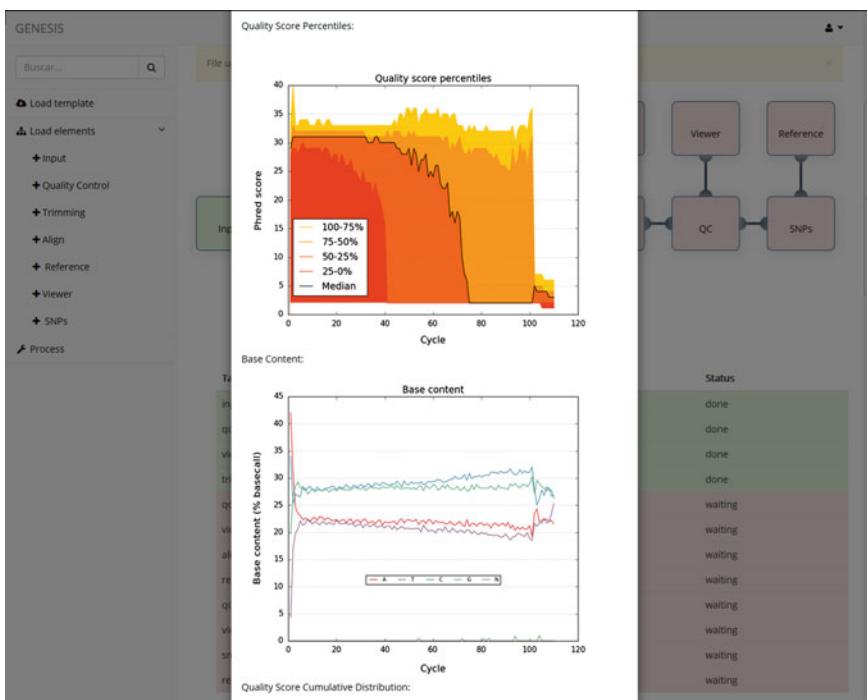
**Fig. 3** Interface of Genesis prototype with an automatic pipeline analysis loaded. Below, real-time task's status system: update of the corresponding block of each task when it finished

(ii) tools menu including tabs for building analysis pipeline (left side vertical menu), and (iii) working area where the different components of the pipeline and resume table are displayed.

To satisfy with the user requirements, User Centered Design (UCD) has been used. Researchers from BioDonostia Health Research Institute have participated actively in the outlook of the prototype (as an example, Fig. 4. shows the quality control online reporter). They provided with recommendations so that the use of the platform results intuitive and really allows the user to save time in the design of the analysis workflow for NGS data.

#### (b) Deployment of framework and web server

Genesis is implemented using Django framework [25], a Python web framework that encourages rapid deployment of web application on a server and particularly suitable for cloud-based applications that include Python libraries. For the deployment, Apache HTTP Server Project has been used. All the databases of the system have been designed using the Django Object-Relational Mapper (ORM).



**Fig. 4** Quality control analysis report outlook

## 5 Conclusions and Future Work

In this paper a cloud-based system for NGS analysis have been successfully developed. The proposed SaaS prototype improves the actual solutions for analysis of NGS data by adding visual features to the interface so that the user can easily design the needed pipeline. User Centered Design approach has been used so that Genesis platform results ergonomic and intuitive, allowing researchers to save time in the implementation and the analysis of their data. Additionally, a first version of an automatic workflow design has been implemented that, in a near future, will provide automatically different analysis pipelines that depends on the target of the experiment (what the user is looking for) and on the sequencer type. In this way, the proposed approach pretends to enable NGS data analysis for a large number of users, without specific IT or bioinformatics knowledge as requirement.

Future work will focus on extending the functionalities of the system by adding further analysis pipelines in order to cover a large spectrum of NGS analysis applications. Additionally, modules to analyze data sets from different omics levels (transcriptomic, proteomic...) will be integrated as well as, interactive interfaces and analytics tools, always with the objective of providing an efficient and easy-to-use solution to biomedical researchers and professionals.

**Acknowledgments** This work was supported by the Provincial Council of Gipuzkoa. The authors would like to express their gratitude to the researchers of the Multiple Sclerosis group of BioDonostia Health Institute for their cooperation.

## References

1. Mardis, E.R.: The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**(3), 133–141 (2008)
2. Quail, M., Smith, M.E., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y.: A tale of three next generation sequencing platforms: comparison of ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics* **13**(1), 341 (2012)
3. Shendure, J., Ji, Hanlee: Next-generation DNA sequencing. *Nat. Biotechnol.* **26**(10), 1135–1145 (2008)
4. Bhuvaneshwar, K., Sulakhe, D., Gauba, R., Rodriguez, A., Madduri, R., Dave, U., Lacinski, L., Foster, I., Gusev, Y., Madhavan, S.: A case study for cloud based high throughput analysis of NGS data using the globus genomics system. *Comput. Struct. Biotechnol. J.* **13**, 64–74 (2015)
5. Thakur, R.S., Bandopadhyay, R., Chaudhary, B., Chatterjee, S.: Now and next-generation sequencing techniques: future of sequence analysis using cloud computing. *Front. Gene* **3** (2012)
6. Nagasaki, H., Mochizuki, T., Kodama, Y., Saruhashi, S., Morizaki, S., Sugawara, H., Ohyanagi, H., Kurata, N., Okubo, K., Takagi, T., Kaminuma, E., Nakamura, Y.: DDBJ read annotation pipeline: A cloud computing-based pipeline for high-throughput analysis of next-generation sequencing data. *DNA Res.* **20**(4), 383–390 (2013)

7. Goecks, J., Nekrutenko, A., Taylor, J.: Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010)
8. Rex, D.E., Ma, J.Q., Toga, A.W.: The LONI pipeline processing environment. *Neuroimage* **19**, 1033–1048 (2003)
9. Hull, D., Wolstencroft, K., Stevens, R., et al.: Taverna: a tool for building and running workflows of services. *Nucleic Acids Res.* **34**, W729–W732 (2006)
10. Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., Krabichler, B., Speicher, M.R., Zschocke, J., Trajanoski, Z.: A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings Bioinform.* **15**(2), 256–278 (2013)
11. Torri, F., Dinov, I.D., Zamanyan, A. et al.: Next generation sequence analysis and computational genomics using graphical pipeline workflows. *Genes* **3**(4), 545–575 (2012)
12. Celery. <http://www.celeryproject.org/>. Accessed 22 Feb 2016
13. RabbitMQ. <https://www.rabbitmq.com/>. Accessed: 22 Feb 2016
14. SQLAlchemy. <http://www.sqlalchemy.org/>. Accessed: 22 Feb 2016
15. Boto. <https://pypi.python.org/pypi/boto>. Accessed: 22 Feb 2016
16. FastqP. <https://pypi.python.org/pypi/fastqp>. Accessed 22 Feb 2016
17. Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L.: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Fenome Biol.* **10**(3), R25 (2009)
18. Li, H.: A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* **27**(21), 2987–2993 (2011)
19. Danecek, P., Auton, A., et al.: The variant call format and VCFtools. *Bioinformatics* **27**(15), 2156–2158 (2011)
20. Subprocess. <https://docs.python.org/2/library/subprocess.html>. Accessed 22 Feb 2016
21. Matplotlib. <http://matplotlib.org/>. Accessed 22 Feb 2016
22. JBrowse. <http://jbrowse.org/>. Accessed 22 Feb 2016
23. Bootstrap. <http://getbootstrap.com/>. Accessed 22 Feb 2016
24. JSPlumb. <https://jsplumbtoolkit.com/>. Accessed 22 Feb 2016
25. Django. <https://www.djangoproject.com/>. Accessed 18 Jan 2016

# Review of Automatic Segmentation Methods of White Matter Lesions on MRI Data

Darya Chyzyk, Manuel Graña and Gerhard Ritter

**Abstract** White matter (WM) lesions are a phenomena perceived in magnetic resonance imaging (MRI) which is prevalent in many different brain pathologies, hence the general interest in automated methods for lesion segmentation (LS). We provide a short review of some commonly used state-of-the-art approaches. The article is focused on the machine learning techniques which researches use to construct semi- and fully-automated tools for LS. In addition, we mention the preprocessing steps, features extraction, LS databases and validation techniques.

## 1 Introduction

Alterations of the myeline covering of the neural sinapses in white matter (WM) producing lesions that disrupt the normal function of the brain are prevalent in many brain pathologies, so automated methods for lesion segmentation (LS) in medical images are of paramount importance. Lesions are most studied in relation with Multiple sclerosis (MS) [10, 14, 15], which affects more than 2.3 million people around the world, so that it is the most widespread disabling neurological condition of young (30–40 years) adults around the world. But white matter lesions appear also in Alzheimer’s Disease, Parkinson’s Sindrome, Mild Cognitive Impairment [7], normal aging, and Arterial Vascular Diseases [1, 2]. Therefore automatic tools for lesion segmentation and measurement may be of great impact in longitudinal studies as well as *in vivo* evaluation of the effect of treatments.

Lesions are visible in magnetic resonance imaging (MRI) modalities with different sensitivities. In Proton Density Weighted Image (PD), T1 weighted, and fluid-attenuated inversion recovery (FLAIR) images lesion appear as hyperintense signal, so that some people refer to lesions as white matter hyperintensities (WMH).

---

D. Chyzyk · M. Graña (✉)  
GIC Research Group of the UPV/EHU, Donostia/San Sebastián, Spain  
e-mail: manuel.grana@ehu.es

D. Chyzyk · G. Ritter  
CISE Department, University of Florida, Gainesville, USA

The design of the image modalities to be captured in a study is of paramount importance for the success of the study.

The task of LS has been attacked from many perspectives [5], from simple thresholding up to semantic analysis based on brain atlases. It is no trivial task for several reasons. From the point of view of thresholding and image filtering, the WMH are a local phenomena, so that similar absolute values can be found in other parts of the healthy brain tissue. Intensity inhomogeneities also affect the interpretation of image intensities, which can vary influenced by the slow frequency inhomogeneity field. From the point of view of machine learning approaches, where LS is tackled as a classification problem, lesions are a very small fraction of the brain volume, leading to strong imbalanced classification problems. The literature is vast, and growing. Due to the artifacts of the image, preprocessing is critical for successful completion of the task. Finally, LS approaches must be endorsed by appropriate validation processes. In this paper we comment on the solutions given in the literature to these issues. A summary encompassing several salient references is provided in Table 1.

## 2 Experimental Databases

The sensitivity of MRI modalities varies also among brain diseases, so that the specific design of the imaging protocols may vary from one to another. The conventional modality combinations for MS studies are T2 and PD modalities, or T2 and FLAIR. The most popular example of single-signal approach is FLAIR. Some studies include Diffusion tensor imaging (DTI). The most extensive study in [1, 2] included T1, inversion recovery (IR), proton density-weighted (PD), T2 and FLAIR. Multimodal studies allow more precise detection of lesion edges.

Most image databases are privative, as public open databases the popular one is the MICCAI's 2008 MS lesion challenge dataset, consisting of 54 subjects with T1, T2 and FLAIR images per subject [10]. Public datasets are very important to foster research in an area, so we will be expecting more datasets made available in the near future. Some examples of privative datasets reported in the literature are:

- ENVISION study [10] (24 elderly subjects)
- The Johns Hopkins Medicine IRB collection [14] (98 MS subjects)
- The work in [1, 2] was done over 19 patients with arterial vascular disease with small (<3 mm), moderate (3–10 mm) and large (>10 mm) lesion
- 10 subjects with MCI and 10 subject with mild dementia [7]
- Hospital Vall D'Hebron, Barcelona (Spain) [15] (70 MS subjects).
- Rotterdam Scan Study. [4] (209 elderly subjects)

**Table 1** Resume table. Pathologies: mild cognitive impairment (MCI), TFM-tissue probability map: Image modalities: *IR* inversion recovery

#	Article	Modality	Disease	Features	Method	Validation
1	[10]	T1, FLAIR	Elderly, MS lesion	Intensity, TPM, coordinates, reference points	filtering, random forest classification, Markov random field	DSI
2	[14]	T1, T2, FLAIR	MS	Unnormalized, Normalized, thresholded, smoothed and moment voxel intensities	Logistic regression, linear discriminant analysis, quadratic discriminant analysis, Gaussian mixture model, SVM, random forest, kNN, neural network, super learner	Receiver operating characteristic (pROC) curve, scaled partial area under the curve (pAUC) as measures of algorithm accuracy. DSC as a measure of agreement amongst the binary segmentations.
					Computational time.	manual segmentation
3	[1, 2]	T1, IR, PD, T2, FLAIR	Arterial vascular disease	Intensity features and spatial information	kNN ( $k = 100$ )	ROC curves, sensitivity and specificity, SI, OF, EF, PSI, POF, PEF
4	[7]	T1, Flair	MCI, mild dementia	Intensities, WM probability, neighborhood, MNI coordinates, 2D Gabor filters (GF)	kNN, SVM	Similarity index
5	[15]	T1, FLAIR	MS	Intensities, WM, GM and CSF probability map	SLS and LST toolbox	Pearson's correlation, ANOVA, post-hoc pairwise significant t-tests with Bonferroni correction
6	[4]	T1, PD	Elderly	Intensity	kNN ( $k = 45$ ), thresholding,	Visual, SI, sensitivity
7	[3]	FLAIR	MS	Intensity	Genetic algorithm, maximum fuzzy entropy	SI, ICC, paired samples t test

Validation measures: *DSI* dice similarity index, *SI* the similarity index, *OF* the overlap fraction, *PSI* the probabilistic similarity index, *POF* probabilistic overlap fraction, *PEF* probabilistic extra fraction compared with other methods, *ICC* Intraclass Correlation Coefficient

### 3 MRI Preprocessing

Any analysis of the MRI image is difficult because of variability of the artefacts of each imaging modality, noise, motion artifacts due to the long imaging time, partial volume effects that difficult the segmentation due to blurred edges. Most frequent preprocessing steps consist in [5, 10]:

- Co-registration of images of different modalities to the anatomical space or MNI standard space. Rigid body transformation registration is applied.
- Removal of non brain tissue from the T1 image (brain extraction tool (BET) by FSL). FLAIR brain extraction is often performed applying the T1 brain mask.

Some optional steps applied by authors:

- Tissue segmentation. For instance, [7] use the WM probability map to generate additional features.
- Radio frequency field inhomogeneity correction is performed on all images using the N3 algorithm [10, 14].
- Intensity normalization among the patients.
- Scaling. Some authors [7] use scaling for all the features, even for MNI coordinates.
- Atlas registration for the identification of likely lesion locations. For instance, [4] pays much attention to a systematic analysis by atlas registration methods. The authors use both single-atlas and multiple-atlas (up to 12 atlases) method with affine registration and non-rigid registration.

The software used for image preprocessing covers a wide collection. The well-known, such as packages for neuroimage processing SPM [11] and FSL are the most used. But some other, like Medical Image Processing Analysis and Visualization (MIPAV), TOADS-CRUISE (<http://www.nitrc.org/projects/toadscruise/>), and Java Image Science Toolkit (JIST) software packages are also used [14].

### 4 Segmentation Methods

The main problems that must be confronted by LS methods are :

- Accurate identification of WMH lesions is difficult due to variability in lesion location, size and shape and anatomical variability between subjects.
- Manual segmentation of MS lesions requires expert knowledge, time consuming and suffers from large intra- and interexpert variability.
- Heterogeneity of lesions and variability in the magnetic resonance (MR) acquisitions parameters, that limits the portability of systems developed with data from one site, and makes difficult to carry multi-site studies requiring homogeneous data.

- Limited number of resources freely available to the community. The range of the open source codes and open datasets available to the community and radiologist tools is small.

Most approaches in the literature use some or several machine learning techniques, which can be combined with preprocessing steps discussed above. Machine learning methods need a precise and appropriate definition of the classification features, so first we will comment on the features discussed in the literature, second we comment on different computational methods.

## 4.1 Feature Space

The basic assumption is that different tissues have different intensities, so it would be possible to carry out the WMH segmentation using only intensity information at each voxel. However, image artifacts and the lack of unicity of the lesion intensities imply that sometimes additional information is used, such as:

- tissue probability maps from the control group or probabilistic atlas [10].
- spatial coordinates [1, 2, 10], which can be MNI coordinates [7].
- Global reference points based contrast (GRCF) [10].
- Spatial neighborhoods which can be entered as raw data [14] or after some kind of filtering, such as the Gabor filters (GF) [7].
- Some postprocessing of the features, such as smoothing or computing spatial moments [14].

For selecting the more discriminant features some authors recommend Adaboost, PCA (using first components with threshold), partition the image into homogeneous regions using the segmentation-by-weighted-aggregations. Other methods use the supervised classification not at the voxel level but at a lesion level: the image is partitioned into homogeneous regions and each region is classified by the method as a lesion or not.

## 4.2 Methods

Most machine learning approaches fall into two categories, supervised and unsupervised approaches. Here we will identify the most frequently used methods of each kind, taking into account that sometimes authors combine several approaches into a pipeline. Pipelines are discussed in the next section.

*Unsupervised methods* Markov Random Field (MRF) can be used in an unsupervised way in order to achieve image correction or regularization based on the minimization of appropriately defined energy function [10], or for outlier detection applying expectation maximization (EM) [6]. Mostly the application of unsupervised approaches

refer to clustering methods such as the fuzzy c-means [3, 12] intended to compress the information in the feature space in order to achieve better classification, or simply to segment the intensity data into regions for further processing. The Otsu's method for the optimal threshold calculation for each image is very specific instance of unsupervised method [7]. Some postprocessing, such as removing non-connected components in the detection image [4] or region growing on a feature map [11], are unsupervised filters. Local histogram analysis [9] performs outlier detection in small neighborhoods

*Supervised methods* Most popular algorithms are classification methods such as random forest (RF) [8, 10, 14], Logistic Regression, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Gaussian Mixture Model, Support Vector Machine [7, 14], k-Nearest Neighbor (kNN) [1, 2, 4, 13, 14], Neural Network, Super Learner [14]. All these methods are very sensitive to the feature space chosen and to the data imbalance, so that correction methods may be needed after the classification. Belief maps [11, 15] also rely of the voxel class label provided by the experts.

A very special class of supervision is the use of the atlas information in the lesion detection [4], which is done after the classification step either as a disambiguation process or as the semantic step which allows to filter out meaningless detections.

## 5 Pipelines

Most approaches are not simple application of one method, they are combinations of several techniques in a computational pipeline starting from the preprocessing of the data.

- [10] uses FLAIR and T1 images, performing a image intensity normalization, classification of image intensity vectors by RF, and correction of edges using MRF based on the probability scores provided by the RF.
- [14] was focused in feature extraction evaluation before classification procedure: voxel intensity and functions of voxel neighborhoods as spatial information. They analyzed 6 type of features: (1) Unnormalized feature vector which contains the observed voxel intensities (after image pre-processing) for a voxel from the T1-w, T2-w and FLAIR volumes; (2) Normalized voxel intensities (three imaging modalities); (3) Thresholded with 85th percentile; (4) the smoothed feature; (5) moments, incorporating spatial information from; (6) Smoothed and Moments.
- [7] compares fully automated supervised methods that learn to identify white matter hyperintensities (WMH). For that they used kNN ( $k = 100$ ) and SVM (non-linear SVM with a radial basis function kernel). They carry out Otsu's thresholding method for the optimal threshold calculation for each image and the stable threshold for the FLAIR images were explored, but did not meet expectations showing poor performance. Furthermore the authors analysed different feature vectors: intensities, WM probability, information about neighborhood, MNI coordinates, 2D Gabor filters as an alternative strategy to extract information about the neighborhood.

- [4] extract WM, GM and CSF segmentations maps using kNN classifiers, and information from brain atlas registration. They have an additional post-processing step by WM fraction calculation, and neighboring connectivity components removal of WM lesion outside of the brain.
- [1, 2] use kNN for fully automated segmentation of WMH. As a result they generate the probability maps of belong to WML, which is thresholded to obtain binary segmentation.
- [14] presented comparison of 9 classification algorithms, and 6 feature extraction methods. As a feature combination they proposed to use normalised, smoothed and moment intensity and additional neighboring information.
- [15] presents the analysis of lesion and tissue automated segmentation on two public available toolkits implemented for the SPM. They include the pre-processing steps based on SPM and different LS strategies. One using T1 for tissue segmentation lesion separation from the WM and thresholding in FLAIR. LST uses T1 images for initial tissue segmentation and T1 and FLAIR for construction a lesion belief map, which was improved later using WM, GM and lesion likelihood information.

## 6 Validation

Validation procedure and tools are of outmost importance [5]. The repository of images most widely employed for validation of LS algorithms applied to MS lesions are those from BrainWeb ([www.bic.mni.mcgill.ca/ServicesBrainWeb](http://www.bic.mni.mcgill.ca/ServicesBrainWeb)), an online database of synthetic MR images. As another option: trained using clinical data and then applied to the BrainWeb images.

Validation encompasses the computation against the ground truth of several standard indices of agreement with the manual segmentation: DICE similarity index [10], precision—recall curves [7], Receiver Operating Characteristic (pROC) curve and scaled partial Area Under the Curve (pAUC) [14]. Some authors use probabilistic similarity index (PSI), probabilistic overlap fraction (POF), probabilistic extra fraction (PEF), maximum similarity index (SI), the overlap fraction (OF), and the extra fraction (EF) [1, 2], because they carry out two types of evaluation: evaluation of binary segmentation and by probabilistic evaluation. In [15] statistical analysis (ANOVA model, post-hoc pairwise significant t-tests with Bonferroni correction) show the differences between methods. To present the correlation in tissue segmentation and lesion volume the Pearson's linear correlation coefficient was used. Some authors [14] include computational time as measure of algorithm competitive quality.

An special case is [4] which mentioned the visual inspection as an authorized method of validation. This type of testing was selected because of the large data of the 209 subjects. The inspection was done by team of two experts. The possible assessment options: good, reasonable, poor for tissue segmentation and no FP or FN, over-, under-segmentation for WM LS.

## 7 Conclusions

This article presents a brief review on the fundamental methodology of MRI brain lesion segmentation. We present the introduction to the main machine learning algorithms commonly used nowadays to localize the brain lesions. The important point mentioned here is to have updated information about new MRI modalities, and possible heir combination to achieve better results. Due to the growing interest to multimodal techniques, it's important to be accurate in preprocessing the data. Co-registration in case of using different MRI modalities, skull-stripping, tissue segmentation, radio frequency field inhomogeneity correction and intensity normalization are the most frequent mention steps. The most popular supervised and unsupervised methods and features have been discussed.

## References

1. Anbeek, P., Vincken, K.L., van Osch, M.J.P., Bisschops, R.H.C., van der Grond, J.: Automatic segmentation of different-sized white matter lesions by voxel probability estimation. *Med. Image Anal.* **8**(3), 205–215 (2004) (Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003)
2. Anbeek, P., Vincken, K.L., van Osch, M.J.P., Bisschops, R.H.C., van der Grond, J.: Probabilistic segmentation of white matter lesions in mri imaging. *NeuroImage* **21**(3), 1037–1044 (2004)
3. Bijar, A., Khayati, R., Penalver-Benavent, A.: Increasing the contrast of the brain mr flair images using fuzzy membership functions and structural similarity indices in order to segment ms lesions. *PLoS ONE* **8**(6), e65469 (2013)
4. de Boer, R., Vrooman, H.A., van der Lijn, F., Vernooij, M.W., Arfan Ikram, M., van der Lugt, A., Breteler, M.M.B., Niessen, W.J.: White matter lesion extension to automatic brain tissue segmentation on mri. *NeuroImage* **45**(4), 1151–1161 (2009)
5. Garcia-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D.L., Collins, D.L.: Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med. Image Anal.* **17**(1), 1–18 (2013)
6. Jain, S., Sima, D.M., Ribbens, A., Cambron, M., Maertens, A., Van Hecke, W., De Mey, J., Barkhof, F., Steenwijk, M.D., Daams, M., Maes, F., Van Huffel, S., Vrenken, H., Smeets, D.: Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *NeuroImage: Clinical* **8**, 367–375 (2015)
7. Klöppel, S., Abdulkadir, A., Hadjidemetriou, S., Issleib, S., Frings, L., Thanh, T., Mader, I., Teipel, S., Hüll, M., Ronneberger, O.: A comparison of different automated methods for the detection of white matter lesions in MRI data. *NeuroImage* **57**(2), 416–422 (2011)
8. Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvador, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., Christensen, Soren, Carey, L.: Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage* **98**, 324–335 (2014)
9. Ramirez, J., Gibson, E., Quddus, A., Lobaugh, N.J., Feinstein, A., Levine, B., Scott, C.J.M., Levy-Cooperman, N., Gao, F.Q., Black, S.E.: Lesion explorer: a comprehensive segmentation and parcellation package to obtain regional volumetrics for subcortical hyperintensities and intracranial tissue. *NeuroImage* **54**(2), 963–973 (2011)
10. Roy, P.K., Bhuiyan, A., Janke, A., Desmond, P.M., Wong, T.Y., Abhayaratna, W.P., Storey, E., Ramamohanarao, K.: Automatic white matter lesion segmentation using contrast enhanced flair intensity and markov random field. *Comput. Med. Imaging Graph.* **45**, 102–111 (2015)

11. Schmidt, P., Gaser, C., Arsic, M., Buck, D., Forschler, A., Berthele, A., Hoshi, M., Ilg, R., Schmid, V.J., Zimmer, C., Hemmer, B., Muhlau, M.: An automated tool for detection of flair-hyperintense white-matter lesions in multiple sclerosis. *NeuroImage* **59**(4), 3774–3783 (2012)
12. Seghier, M.L., Ramlackhansingh, A., Crinion, J., Leff, A.P., Price, C.J.: Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *NeuroImage* **41**(4), 1253–1266 (2008)
13. Steenwijk, M.D., Pouwels, P.J.W., Daams, M., van Dalen, J.W., Caan, M.W.A., Richard, E., Barkhof, F., Vrenken, H.: Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (knn-ttps). *NeuroImage: Clinical* **3**, 462–469 (2013)
14. Sweeney, E.M., Vogelstein, J.T., Cuzzocreo, J.L., Calabresi, P.A., Reich, D.S., Crainiceanu, C.M., Shinohara, R.T.: A comparison of supervised machine learning algorithms and feature vectors for ms lesion segmentation using multimodal structural mri. *PLoS ONE* **9**(4), e95753 (2014)
15. Valverde, S., Oliver, A., Roura, E., Pareto, D., Vilanova, J.C., Ramio-Torrenta, L., Sastre-Garriga, J., Montalban, X., Rovira, A., Llado, X.: Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling. *NeuroImage: Clinical* (2015)

# Hygehos Ontology for Electronic Health Records

Naiara Muro, Eider Sanchez, Manuel Graña, Eduardo Carrasco,  
Fran Manzano, Jose María Susperregi, Agustín Agirre  
and Jesús Gómez

**Abstract** During the last years a high effort on standardization of Electronic Health Records has been made. Standards ISO EN 13606 and OpenEHR with their dual approach have promoted semantic interoperability in the real clinical practice. Recently, the focus has been set on the extraction of knowledge from the clinical information stored in EHR, but current approaches based on archetypes do not provide a complete solution regarding content structuration limitation. In this paper we propose an ontology for Hygehos Electronic Health Records (EHR), that we call the Hygehos Ontology. The introduction of such ontology on the EHR system will facilitate the development of reasoning and knowledge extraction tools over the stored clinical information. In our approach, we first align the Hygehos EHR to the dual model of OpenEHR and generate the corresponding archetypes for every part of the system. Secondly, we formalize a methodology for structuring the clinical contents of Hygehos EHR into the Hygehos Ontology.

**Keywords** EHR • ADL archetype • OWL ontology • Hygehos

---

The original version of this chapter was revised: Affected figure has been replaced. The erratum to this chapter is available at [10.1007/978-3-319-39687-3\\_33](https://doi.org/10.1007/978-3-319-39687-3_33)

---

N. Muro (✉) · E. Sanchez · E. Carrasco  
Vicomtech-IK4, San Sebastian, Spain  
e-mail: nmuro@vicomtech.org

N. Muro · M. Graña  
Computational Intelligence Group, University of the Basque Country UPV/EHU,  
San Sebastian, Spain

E. Sanchez · E. Carrasco  
Biodonostia Health Research Institute, San Sebastian, Spain

F. Manzano  
Igarle, San Sebastian, Spain

J.M. Susperregi · A. Agirre · J. Gómez  
La Asunción Clinic, Tolosa, Spain

## 1 Introduction

Interoperability and clinical data processing play an important role in biomedical research through a variety of applications linked to hospitals' databases. Different approaches have been applied in the search of interoperability between heterogeneous information systems [1].

Standardized Electronic Health Records have allowed the communication and interpretation of this data within a wide variety of medical centers [2]. Its dual structure composed of a reference information model and an archetype model has allowed separation between clinical content and its structuration, in order to promote this interoperability [3]. However, the archetypes responsible for the definition of clinical concepts are not able to support complex reasoning and knowledge discovery requirements [4]. A more complete structuration of the information is needed to overcome this lack in order to be able to process semantic queries.

Ontologies are used primarily as a source of vocabulary standardization and integration, and many applications can be benefited of using them for computable knowledge extraction. Transforming the current archetypes into OWL (Web Ontology Language) classes and extending them into an ontology would overcome the current lack of ADL (Archetype Definition Language) archetypes when introducing the obtained reasoning conclusions from the content in their own definition [5, 6].

In this paper we propose an ontology for Electronic Health Records (EHR), that we call the Hygehos Ontology. The introduction of such ontology on the EHR system will allow the application of reasoning and knowledge extraction tools over the stored clinical information. We based our approach on the Hygehos EHR [7]. We first align the Hygehos EHR with the dual model of the OpenEHR standard and generated the corresponding archetypes for every part of the system. Then, we formalize the methodology that we have followed for structuring the clinical contents of the Hygehos EHR into the Hygehos Ontology.

This paper is structured as follows: in Sect. 2 we present a brief state of the art of relevant concepts and approaches; in Sect. 3 we describe the methodology followed by the Hygehos system; in Sect. 4 such methodology is formalized into the Hygehos Ontology; in Sect. 5 we present the mapping between the Hygehos Ontology with the dual modelling of Hygehos to OpenEHR, and finally in Sect. 6 we summarize conclusions and future work.

## 2 Background Concepts

This section describes some relevant technologies of our approach, such as ontologies and EHR. Also, we briefly introduce the Hygehos EHR, over which our work has been developed.

## 2.1 *Ontologies*

Following Gruber's definition [8] ontologies are formal specifications of a conceptualization. Ontologies allow the descriptive representation of rich and complex knowledge about concepts and their relations. They determine semantic identifiers and formal descriptions representing the classes of entities of a specific domain [9]. Ontologies have been applied in the literature to the clinical domain in different approaches and applications, such as (i) the agreement about the knowledge model of a domain [10], (ii) clinical decision support systems [11] and recently also for EHR enhancement [12, 13]. In particular the work of Beale and Heard [12] describe an interoperable and safely computable clinical information model based on an ontological analysis of the process of clinical care delivery, seen as a scientific problem-solving process. The works of [13] address the semantic interoperability of two EHR standards (OpenEHR and ISO EN 13606) by a solution capable of transforming OpenEHR archetypes into ISO EN 13606 and vice versa that combines Semantic Web and Model-driven Engineering technologies.

## 2.2 *EHR*

EHR are defined as structured clinical data repositories. Provided that certain minimum requirements are met regarding ubiquity, they allow accessibility from anywhere at any time [14]. Their main features are: (i) to ensure readable, interpretable and persistent information record, (ii) to provide a unequivocal and unique identification for each patient, (iii) to allow interoperability between different healthcare centers as well as between different departments within the same center, (iv) to apply a standardization of the recorded data in the EHR to make interoperability possible, and (v) to facilitate the use and access to all patients' EHR and aid with the visualization and processing of the data. All this must be applied taking into account the security and privacy policy imposed by law to clinical data, and ensuring the authenticity of all documents filed in the EHR, through a signature of the responsible [15]. The multiple features described before have made its implementation in medical centers very favorable and widespread.

Over the last decade different EHR standards had been developed for EHR. The most extended ones are (i) HL7, (ii) ISO EN13606 and (iii) OpenEHR. Although each standard brings a particular feature compared to others, all agree on a dual model structure, consisting on a reference model (RM) and an Archetype Model (AM). The RM supports information within a structure, based on well-established concepts independent from knowledge. It represents the characteristics of the general components and their organization. AM defines and models concepts of clinical knowledge following the structure and constraints imposed by the RM. The combination of both models in a single frame provides of structure and semantic interpretation to the content stored in the EHR.

Archetypes are plain text files written following a syntax called ADL. Their goal is to represent particular clinical concepts without taking into account their information representation structure, giving only their clinical definition. The dual model approach was implemented just for that reason: to separate the clinical concept from the information model that embodies it. In that way particular clinical concepts are represented as a set of constraints on the generic model of information [14].

As we reported in our previous work [16], some challenges concerning knowledge extraction processes from the EHR that follow from the dual model are still unsolved. For that reason, we proposed in [16] an extension of the dual model to a triple model approach. This model extension proffers a new structure, which instead of relying directly on the timescale defined by the facts (clinical documents), was actually based on a scale defined by time dependent Decisional Events (DE). In this work, we present an ontology that can serve as a mapper between the document-based approach to a DE-based approach, allowing application of reasoning and knowledge extraction tools into EHR.

### 3 Hygehos Methodology

Hygehos is a proprietary Electronic Health Record developed during the last 15 years by the Spanish IT companies Igarle and STT and the clinical team of the La Asunción clinic.<sup>1</sup> The system covers hospital-side information system, but also contains a module, called Hygehos Home, for remote monitoring of patients and patient-doctor telecommunication services [17].

Hygehos follows a generic approach for the acquisition of the clinical information of patients. This fact allows the direct implantation of such system into almost every medical center and hospital, independently to the specialization level or type. Currently it is running in more than 15 hospitals and medical centers in Spain, each of them oriented to a different specialization, for instance a monographic center in oncology and a primary care center. Such generic approach, the vision and the methodology for EHR followed by Hygehos are presented next.

Hygehos collects all the information concerning the status of the patient and classifies it into three types: (i) Permanent Information, (ii) Episodic Information and (iii) Evolution Information. The Permanent Information contains relevant patient information of general interest for any health professional in charge of that patient as clinical background, vaccinations, living will and social history, amongst others. It has to be always available for query submissions and it lasts over time, increasing its contents. It will be applied in any clinical procedure regardless of its origin source.

The Episodic Information is defined as information limited in time, i.e. isolated facts. The measurement unit of the episodic information is an episode. Within it, we

---

<sup>1</sup>Collaboration between La Asunción Clinic (<http://www.clinicadelaasuncion.com/>), Igarle (<http://www.igarle.es/>), and STT (<http://www.stt-systems.com/>).

can distinguish three levels of information: (i) Emergencies (e.g. home-based, hospital-based), (ii) hospitalization (e.g. acute, subacute, home hospitalization) and (iii) outpatient (e.g. consultations, check-ups). For each episode different medical acts are made: (i) preventives (e.g. vaccinations, check-ups), (ii) diagnosis (e.g. radiography, colonoscopy, cytology, spirometry, blood analysis), (iii) therapies (e.g. surgery, rehabilitation), (iv) research (e.g. clinical trials) and (v) Aesthetic (e.g. Botulinum Toxin), amongst others. Depending on the type of each clinical procedure, some clinical documents are associated to it. For some of these clinical documents, there are some limits regarding minimum contents definition by law.

Hygehos makes a full scan of the treated person, not only a punctual study, so each episode is not understood as an isolated event but as different phases of clinical processes suffered by the patient. We can understand episodes as isolated or related to a process. As an example, pregnancy could be considered as a clinical process and different episodes will be associated to it (e.g. a gestational diabetes caused by hormonal change in pregnancy). For other cases in which the episode is related to a specific patient's condition or a particular clinical situation, such as being chronic heart failure patient, the different episodes relate in a particular way to the clinical processes. The definition of these relationships allows the extraction of global conclusions for personalized patient care.

The Evolutive Information is generated with a temporal discontinuity. It is composed by documents with previous annotations that are permanent and unalterable but that allow the adhesion of new comments or notes resulting from care processes. Combined together, they form a unit in content terms (clinical course, evolutionary notes, active treatments, etc.).

## 4 Hygehos Ontology

We have formalized the Hygehos Methodology into an ontology, which we call the Hygehos Ontology (depicted in Fig. 1). It represents the domain model of an Electronic Health Record, all the clinical information contained, as well as the rest of the elements that interact with it.

The main class **ElectronicHealthRecord** is related to class **ClinicalInformation** by the object type property *ehrContains*. **ClinicalInformation** represents every clinical data, which can be of different types, contained in the EHR. A clinical information will always correspond to a patient treated by clinical center's health personal. Class **Patient** is related to **ClinicalInformation** by the object type property *correspondsTo*. Class **Patient** has a data type property *patientId*. Class **HealthPersonal** is related to **ClinicalInformation** by the object type property *hasResponsible*. Class **HealthPersonal** has a data type property *healthPersonalId*. Each health personal is part of a medical unit of the hospital or center, according to their specialization. Class **ClinicalInformation** is related to class **MedicalUnit** by the object type property *isFrom*. Class **MedicalUnit** has a data type property *unitName*.



**Fig. 1** Hygehos ontology

Class **ClinicalInformation** has three different disjoint subclasses: (i) **PermanentInformation**, (ii) **EvolutiveInformation** and (iii) **EpisodicInformation**. In turn, class **PermanentInformation** has six disjoint subclasses: **Allergies**, **Vaccination**, **WorkHistory**, **PersonalHistory**, **FamilyHistory** and **AdministrativeInformation**. Each of these classes contains different data type properties that will describe the different data stored for each case (e.g. class **Allergies** contains a data type property *allergyName*). In order to keep the information clear, Fig. 1 does

not show such properties, and represents the set of data type properties corresponding to each of the last-level-subclasses as a unique empty block.

The set of data type properties corresponding to each of the last-level-subclasses of the Hygehos Ontology are mapped with the contents of archetypes that we implemented for the Hygehos system, as presented in Sect. 5. The same occurs with the rest of the last-level-subclasses of the classes **EvolutiveInformation** and **EpisodicInformation**.

Class **EvolutiveInformation** has two disjoint subclasses (last-level, and thus each of them has associated a set of data type properties): **ClinicalEvolution** and **ActiveMedicalOrders**, which includes the active treatments at the moment but also the ones applied before in order to evaluate the therapeutic indications taken into account. These therapeutic indications covers widespread medical issues, not only pharmacological.

Class **EpisodicInformation** is related to a class **Process** by the object type property *containsProcess*. Class **Process** is related to class **Episode** by the object type property *containsEpisode*. Class **Episode** is related to class **Procedure** by the object type property *containsProcedure*. Class **Procedure** is related to class **ClinicalDocument** by the object type property *containsClinicalDocument*. **ClinicalDocument** has seventeen disjoint subclasses (last-level, and thus each of them has associated a set of data type properties): **ClinicalStatisticalReport**, **EntryAuthorization**, **InformedConsent**, **AnamnesisAndPhysicalExploration**, **Evolution**, **MedicalOrder**, **ComplementaryExplorationReport**, **ReferralReport**, **SurgeryReport**, **Anesthesia**, **BirthReport**, **NursingCarePlan**, **PharmacologicalTreatment**, **VitalSignGraph**, **DischargeClinicalReport**, **NecropsyReport** and **Urgency**. This object will have as many subclasses as possible care events may occur in the present or in the future.

## 5 Knowledge Exploitation from Hygehos

The overall objective of this work is to include reasoning tools in the Hygehos System to extract more advanced and complex conclusions from the stored information into the clinical history. In order to put to work the generated Hygehos Ontology, first we have standardized the clinical history following the dual model of the openEHR standard. We created the different archetypes of the system, 25 in total. Each archetype is mapped to a class of the ontology, and each element of the archetype to a data type property (whose domain class is the one with which the archetype is mapped).

To illustrate the approach we present the archetype for the clinical document of anamnesis and physical exploration in Fig. 2. The elements defined are (i) registration date (data type DV\_DATE\_TIME); (ii) registration responsible professional (data type DV\_TEXT) and corresponding medical unit (data type DV\_TEXT); (iii) matter of the urgency (data type DV\_TEXT); (iv) patient's background or medical history, with its type (data type DV\_TEXT), date (data type DV\_DATE\_TIME) and

```

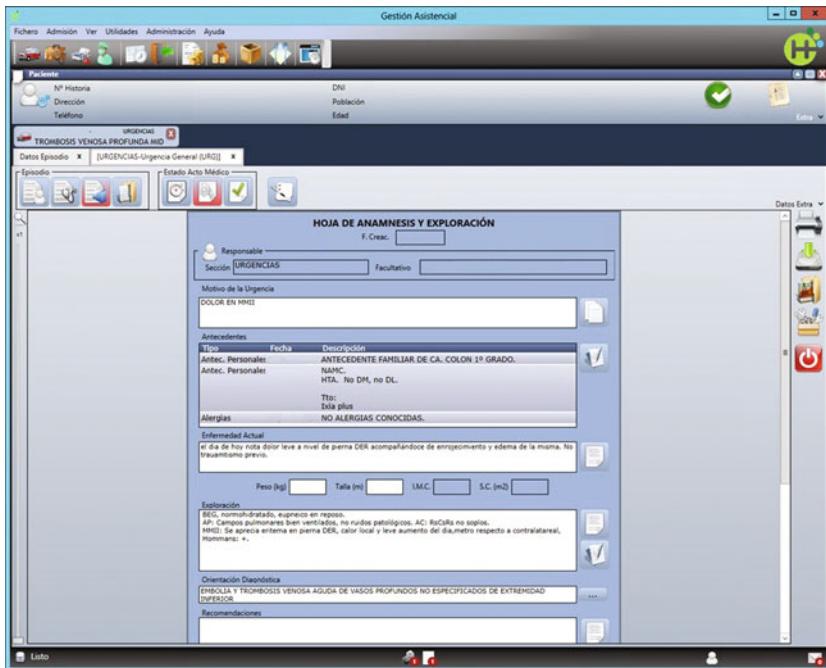
definition
OBSERVATION[at0001] matches {-- Anamnesis and Exploration Document
ELEMENT[at0002] occurrences matches {0..1} matches { -- Date/Time
    value matches {DV_DATE_TIME matches {value matches {yyyy-mm-dd}}}}
CLUSTER[at0003] occurrences matches {0..*} matches { -- Responsible
    items cardinality matches {0..*; unordered} matches {
        ELEMENT[at0003.1] occurrences matches {0..1} matches { -- Section
            value matches {DV_TEXT matches {*}}}
        ELEMENT[at0003.2] occurrences matches {0..1} matches { -- Practitioner
            value matches {DV_TEXT matches {*}}}}
ELEMENT[at0004] occurrences matches {0..1} matches { -- Matter Of Urgency
    value matches {DV_TEXT matches {*}}}
CLUSTER[at0005] occurrences matches {0..*} matches { -- Background/ Medical History
    items cardinality matches {0..*; unordered} matches {
        ELEMENT[at0005.1] occurrences matches {0..1} matches { -- Type
            value matches {DV_TEXT matches {*}}}
        ELEMENT[at0005.2] occurrences matches {0..1} matches { -- Date
            value matches {DV_DATE_TIME matches {value matches {yyyy-mm-dd}}}}
        ELEMENT[at0005.3] occurrences matches {0..1} matches { -- Description
            value matches {DV_TEXT matches {*}}}}
ELEMENT[at0006] occurrences matches {0..1} matches { -- Current illness
    value matches {DV_TEXT matches {*}}}
CLUSTER[at0007] occurrences matches {0..*} matches { -- patient biometrical data
    items cardinality matches {0..*; unordered} matches {
        ELEMENT[at0007.1] matches {-- weight
            name matches {CODED_TEXT matches {code matches {[ac0007.1]} -- weight}}
            value matches {QUANTITY matches {value matches {0..800}units matches {"kg"}}}}
        ELEMENT[at0007.2] matches {-- size
            name matches {CODED_TEXT matches {code matches {[ac0007.2]} -- size}}
            value matches {QUANTITY matches {value matches {0..5}units matches {"m"}}}}
        ELEMENT[at0007.3] matches {-- BMI: Body Mass Index
            name matches {CODED_TEXT matches {code matches {[ac0007.3]} -- BMI}}
            value matches {QUANTITY matches {value matches {0..100}}}}
        ELEMENT[at0007.4] matches {-- body surface
            name matches {CODED_TEXT matches {code matches {[ac0007.4]} -- body surface}}
            value matches {QUANTITY matches {value matches {0..3}units matches {"m^2"}}}}}}
ELEMENT[at0008] occurrences matches {0..1} matches { -- Exploration
    value matches {DV_TEXT matches {*}}}
ELEMENT[at0009] occurrences matches {0..1} matches { -- Diagnosis
    value matches {DV_TEXT matches {*}}}
ELEMENT[at0010] occurrences matches {0..1} matches { -- Recommendations
    value matches {DV_TEXT matches {*}}}}

```

**Fig. 2** Archetype for anamnesis and physical exploration

description (data type DV\_TEXT); (v) current illness for which medical care is requested (data type DV\_TEXT); (vi) patient's biometric data, with the weight (data type CODED\_TEXT), size (data type CODED\_TEXT), body mass index (data type CODED\_TEXT), and body surface (data type CODED\_TEXT); (vii) description of the exploration (data type DV\_TEXT); (viii) diagnosis (data type DV\_TEXT), and (ix) the final recommendations (data type DV\_TEXT). In Fig. 3 the graphical user interface corresponding to the archetype of Fig. 2 is shown. Figure 4 depicts the part of the archetype where the link of each element with the corresponding class or properties of the Hygehos Ontology is shown.

Having the archetype mapped to the ontology allows to reuse of the information contained to generate new conclusions and extract knowledge from the EHR. In our previous work [16] we proposed a new decisional model in which the information contained in the EHR would be represented based on each of the decisions made on a patient, rather than based on the documents and reports generated during each episode. Each decisional event in the new model describes the context in which such decision was made: (i) patient data, (ii) different criteria followed, such as clinical guidelines and protocols, (iii) the objective of the decision, such as the total recovery of the patient or the fastest partial recovery possible, (iv) the final decision value, for instance of the treatment prescribed to the patient, and (v) the result of the decision, in terms of success or failure achieved with respect to the objective. The analysis of such context allows the extraction of low-level conclusions about the patient, their symptoms, the treatments that could succeed and the ones that have had high failure rates in the past [18, 19].



**Fig. 3** User interface for anamnesis and physical exploration

In order to have patient information represented in such a decisional model, a mapping between the current approach (document-based) and the decisional model is needed. The Hygehos Ontology is the formalization of the current document-based model, which is a starting point for such mapping process. We are currently working on the mapping of the Hygehos Ontology into a decisional model.

## 6 Conclusions and Future Work

In this work we have built an ontology, called the Hygehos Ontology, which represents logically the generic methodology followed by the Hygehos EHR system. Our approach aims to provide knowledge extraction capabilities to EHR. The ontology we present in this paper formalizes the document-based approach followed by dual-model-EHR. In this sense, it can be used as an starting point to map such model into a decision based approach followed by the triple we proposed in our previous work [16] and for which we already proposed knowledge extraction and analysis tools [18, 19].

On the other side, we also mapped Hygehos to OpenEHR and provided the corresponding archetypes. In this way our work also allows to separate the

graphical interfaces from the functionality and the contents, by allowing the generation of interfaces directly from archetypes. Such archetypes can be modeled directly by medical experts that do not have software technical expertise by using open archetype authoring tools provided by the OpenEHR community. In this way usability and user experience of Hygehos is improved, and personalized products can be provided to each customer (and their corporate image) with a small effort.

As future work, we will build the knowledge extraction tool based on the Hygehos Ontology. We will also build a tool that generates graphical interfaces directly from the definition of archetypes.

**Acknowledgments** This research was partially funded by the Basque Business Development Agency (SPRI), dependent on the Basque Government, under the grant GAITEK2015-SemanHis.

## References

1. Bellazi, R., Diomidous, M., Sarkar, I.N., Takabayashi, K., Ziegler, A., McCray, A.T.: Data analysis and data mining: current issues in biomedical informatics. *Methods Inf. Med.* **50**, 536–544 (2011)
2. Jha, A.K., DesRoches, C.M., Campbell, E.G., Donelan, K., Rao, S.R., Ferris, T.G., Shields, A., Rosenbaum, S., Blumenthal, D.: Use of electronic health records in U.S. hospitals. *N. Eng. J. Med.* **360**(16), 1628–1638 (2009)
3. Wollersheim, D., Anny, S., Wenny, R.: Archetype-based electronic health records: a literature review and evaluation of their applicability of health data interoperability and access. *Health Inform. Manage. J.* **38**(2), 7–17 (2009)
4. Iqbal, A.M., Shepherd, M., Abibi S.S.R.: An ontology-based electronic medical record for chronic disease management. In: 44th Hawaii International Conference on System Science (HICSS) (2011)
5. Pathak, J., Kiefer, R.C., Chute, C.G.: Using semantic web technologies for cohort identification from electronic health records for clinical research. *AMIA JT Summits Transl Sci Proc.* pp. 10–19 (2012)
6. Lezcano, L., Sicilia, M.A., Rodríguez-Solano, C.: Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. *J. Biomed. Inform.* **44**(2), 343–353 (2010)
7. Carrasco, E., Sanchez, E., Artetxe, A., Toro, C., Graña, M., Guijarro, F., Susperregui, J.M., Aguirre, A.: Hygehos home: an innovative remote follow-up system for chronic patients. In: *Studies in Health Technology and Informatics: Innovation in Medicine and Healthcare 2014*, vol. 207, pp. 261–270, IOS Press (2014)
8. Gruber, T.: What is an Ontology (1993). <http://www-ksl.stanford.edu/kst/whatis-an-ontology.html>. Accessed 07 Sept 2004
9. Martinez-Costa, C., Menarguez-Tortosa, M., Fernandez-Breis, J.T., Maldonado, J.A.: A model-driven approach for representing clinical archetypes for semantic web environments. *J. Biomed. Inform.* **42**(1), 150–164 (2009)
10. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? *Int. J. Hum Comput Stud.* **43**(5), 907–928 (1995)
11. Farion, K., Michalowski, W., Wilk, S., O'Sullivan, D.M., Rubin, S., Weiss, D.: Clinical decision support system for point of care use: ontology driven design and software implementation. *Methods Inf. Med.* **48**(4), 381–390 (2009)
12. Beale, T., Heard, S.: An ontology-based model of clinical information. In: *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems*, p. 760. IOS Press (2007)

13. Martínez-Costa, C., Menárguez-Tortosa, M., Fernández-Breis, J.T.: An approach for the semantic interoperability of ISO EN 13606 and OpenEHR archetypes. *J. Biomed. Inform.* **43**(5), 736–746 (2010). Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? *Int. J. Human-Comput. Stud.* **43**(5), 907–928 (1995)
14. Garde, S., Knaup, P., Hovenga, E.J.S., Heard, S.: Towards semantic interoperability for electronic health records: domain knowledge governance for OpenEHR archetypes. *Methods Inf. Med.* **46**(3), 332–343 (2007)
15. Hayrinne, K., Saranto, K., Nykanen, P.: Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *Int. J. Med. Informatics* **77**(5), 291–304 (2007)
16. Sanchez, E., Toro, C., Graña, M.: Integrating electronic health records in clinical decision support systems. *Innov. Med. Healthcare*, 407–416 (2016)
17. Blobel, B., Kalra, D., Koehn, M., Lunn, M., Pharow, P., Ruotsalainen, P., Schulz, S., Smith, B.: The role of ontologies for sustainable, semantically interoperable and trustworthy EHR solutions. In: European Federation for Medical Informatics (2009)
18. Sanchez, E., Toro, C., Artetxe A., Graña, M., Sanin, C., Szczerbicki, E., Carrasco, E., Guijarro, F.: Bridging challenges of clinical decision support systems with a semantic approach. A case study on breast cancer. *Pattern Recogn. Lett.* **34**(14), 1758–1768 (2013)
19. Sanchez, E., Peng, W., Toro, C., Sanin, C., Graña, M., Szczerbicki, E., Carrasco, E., Guijarro, F., Brualla, L.: Decisional DNA for modeling and reuse of experiential clinical assessments in breast cancer diagnosis and treatment. *Neurocomputing* **146**, 308–318 (2014)

# Views on Electronic Health Record

Manuel Graña and Oier Echaniz

**Abstract** The Electronic Health Record (EHR) is the central information object for healthcare and medical related industries. However, it has been given little academic attention per se, because it is always embedded in the information system of the hospitals. This paper consider several aspects of EHR management systems: security and privacy, data mining, design of decision support systems, acceptance by users and producers of health resources, and system implementation.

## 1 Introduction

The revolution in the information and communication technologies (ICT) has brought a new paradigm to the health sciences and industries. The definition and implementation of Electronic Health Records (EHR) systems and related standards is the cornerstone for the advent of electronic Health Information Exchange (HIE) systems and the full development of Health Information Technologies (HIT), including Health Information Mining (HIM). A good example of the state of affairs is the diversity of names that are present in the literature for the EHR concept under different points of view, such as Electronic Patient Record (EPR), Electronic Medical Record (EMR), Personal Health Record (PHR). The latter is directly related to the empowerment of the individual as the manager of its own health [40]. There are even propositions of family health record systems [5]. On top of them, clinical information models (CIM) are proposed as the next layer [35], guiding information exchange, interoperability, and extraction of value by data mining procedures. A search in sciencedirect and iee-explore with the term “Electronic Health Records” produced over 1.500 references, which is an indication of how the field is thriving. In this review we have focused on the last three years, adding older references when found relevant.

---

M. Graña (✉) · O. Echaniz  
Computational Intelligence Group of the University of the Basque Country,  
San Sebastian, Spain  
e-mail: manuel.grana@ehu.es

The EHR contain retrospective, actual, and prospective information of the patient, encompassing encounters with healthcare providers, medications, history of operations and hospitalizations, laboratory and signal/imaging results, past diagnostic and follow up, and even video recording of patient-provider interactions [43]. It has been also proposed EHR as a key instrument to support public health [27], which implies extending its contents to non-medical environmental information, and providing generalized protocols and processing pipelines make the EHR information available. The HIE connects primary care, hospitals, pharmacies and laboratories to the healthcare information management network, so that the EHR may play a central role in communications between health institutions and companies, as well as inside smart hospitals [33]. Therefore, definition of EHR may involve negotiation processes among diverse stakeholders [39] such as software companies, local service providers, and regulatory agencies.

In this paper we gather several perspectives related to EHR design starting from a discussion of the driving forces for EHR design in Sect. 2, and some legal and ethical issues in Sect. 3. More to the point, Sect. 4 presents the implementation issues with an emphasis on the interoperability of systems and the growing trend towards open source approaches leading to free shared implementations. Section 5 discusses the most important issues on information infrastructures, such as interface design, cloud computing and related security issues such as anonymity. Finally, Sect. 6 gives some conclusions targeting the user acceptance issues.

## 2 Motivation of EHR Design

The core motivation for the design and implementation of EHR and related systems is the improvement of care quality. Access to updated sound (error free and up to date) patient information by the care provider would allow better diagnostic, treatment decisions, and follow up, including tracking medical errors. Patients would also be more able to move between care providers, looking for the best health service, or to carry health self management with increased support from monitoring devices and social support, so that society is steadily moving towards the Personally Controlled EHR (PCEHR) [1, 25] for the management of lifelong health information. In the limit this patient mobility must deal with cross-boundary issues in transnational use of data [18]. In close relation, is the improvement of administration processes, from admission to billing.

The EHR implementations alone were estimated at \$15.5 billion in 2010 and it is projected to grow to \$19.7 billion [38]. The current global mobile health market was valued at \$6.6 billion in 2013 [2]. Besides, the pharmaceutical industries, as well as other health related industries, are interested in EHR data mining [14, 26] looking for innovations, assessment of the effect of drugs, identification of target health markets, and real data for the development of medicines [4]. Also, government agencies are interested in expense prediction in order to allocate resources [56]. All of them need consistent, updated, anonymous, and error free data.

### 3 Legal, Political and Ethical Issues

Security and privacy are primary concerns with strong legal and ethical implications [16]. The key technological elements related to privacy are how to grant access to the data, which access models are applied, and the anonymization of the record. The security technological requirements are related to data encryption and certification as well as identity authentication to reach data access. With the advent of smartphones new challenges are confronted. It is reported that 80 % of health care personnel use their smartphones for work related tasks [6], so that the system becomes a bring your own device (BYOD) environment, where a whole ecology of applications may be coexisting in the devices used for EHR data visualization and manipulation, opening the door to many unintentional but malicious threats. New policies to minimize these threats include installing approved security applications in the personal devices and other intrusive means. Solutions for security enhancement in mobile health ecosystems [48] contemplate loss of connectivity or delays, as well as theft and device sharing. Hence, offline authentication must be added, as well as secure storage by encryption with authenticated keys. Such ecosystems can take the form of mobile health social networks, i.e. among patients suffering the same disease, [57] offering benefits such as opportunistic computing and communication sharing communication links and resources, as well as having social comfort. These social networks require new cryptographic security strategies as well as trust assessment methods to avoid attacks at a diversity of levels, from the body area network to the communication grid. The advent of pervasive computing and the internet of things in the health care systems ecology [52] requires new distributed authentication protocols, which must be robust to power failures.

## 4 System Implementation

There are already a number of standards [35] for the definition of EHR and communication and exploitation systems. Base standards are the EN ISO 13606 Reference Model and HL7 Clinical Document Architecture. Specifications for clinical information representation within EHRs include: ISO EN 13606 archetypes, openEHR archetypes, HL7 Clinical Document Architecture, HL7 Templates, and Detailed Clinical Models (DCM).

### 4.1 Interoperability

The interoperability between systems faces a myriad of standards that are more or less incompatible. In some specific domains, such as orthodontics [32], ad hoc solutions to allow communication between mainstream software were developed.

However, the need for some kind of universal EHR language (UEL) and query system able to translate between specific and universal languages still remains unanswered. A quantum UEL is proposed by [43, 44], which uses for reasoning the new formalism of Hyperbolic Dirac Nets (HDN) [42] imported from quantum mechanics. They explain translation between this formalism and other established like the HL7 and the VISTA specifications. From the point of view of CDSS, transformation between HDN and the conventional Bayesian Nets prove equivalent reasoning power, and the latter are a subset of the former [42]. Another approach tries to harmonize the views of two mainstream standards for communication between personal health devices and base computers [53] aiming to create an ecosystem for new PHR business opportunities. A cloud based open platform for development of PHR apps is proposed in [54] provides interoperability and data protection by using virtual machines as sandboxes.

## 4.2 Open Source and Free Software

Open source allows transparent verification of systems, on the other hand, vendors enforcing confidentiality of private code problems [28] for marketing reasons, hinders the insertion of EHR for lack of trust. Open source and public free software solutions are becoming a critical substrate for the development of new systems. Open source have been widely used for medical image visualization [41]. The main advantage is economical, reducing the high cost of medical solutions. A system of EHR for ambulatory care in Canada [45] called OSCAR has been deployed successfully with good acceptance of care providers and public. The inconveniences of open source solutions are lack of standards, interoperability, and technical support. Another example of large deployment is the VLER [7] which was built upon public gateway software for secure communication over the web. Most tools for Big Data are open source, which hinders the widespread implementation and access in the medical domain which requires straightforward out-of-the-box solutions with comfortable interfaces. However, institutions like the NIH recognize the need to advance in this direction establishing the Big Data to Knowledge infrastructure program providing a shared computational environment (e.g. data standards, ontologies, data catalogues, virtualized cloud computing).

## 5 Information Processing Infrastructure

Since the early designs the need for data mobility, i.e. by smartcards before the Internet full deployment [36], and the security issues that it raises, has been a prime concern. Patient data needs to travel with him in order to provide useful and timely support to the treatment, however it requires also authenticity certification and

privacy [8]. Concerns of security and privacy [29] encompass risks from the user, external attacks, and unethical data mining from companies.

### 5.1 *Cloud Computing*

EHR storage and processing in the cloud promises anywhere, anytime access to critical data. However, privacy and security slip from the hands of the care providers and the patient. In the framework of globally distributed cloud services, provides a traffic shaping algorithm to overcome traffic analysis attacks, and a resource distribution ensuring minimum delay and queue stability. The risk of third party intrusion is very high in the cloud. A encryption approaches [13, 30] allow proper data access control by the authorized users in a context of multiple access levels, providing confidentiality, authenticity, unforgeability, anonymity and collusion resistance.

Secure communication is a priority requirement for EHR pipeline. Secure communication over public Internet of laboratory results as HL7 messages using a combination of off-the-shelf secure tools through a DIRECT gateway has been demonstrated [50].

### 5.2 *Anonymization, De-Identification*

The NIH classifies research involving human as Human Subjects Research (HSR) or Not Human Subjects Research (NHSR), the NHSR having much less administrative oversight and can be more effectively exploited. De-identification (aka anonymization) is the key process for the effective exploitation of EHR data. Private threads relate to direct identifiers, quasi-identifiers, and sensitive attributes [19] which can be subject of identity, membership or attribute disclosures. Privacy models, such as k-Anonymity, allow to design algorithms and to assess the privacy risks of already anonymized data. Some processes achieve anonymization by statistical analysis, removing high entropy variables [11]. Besides structured information in fields, EHR may also contain free form text, which can be cleared of identity information by use of basic syntactic analysis [23] to some extent. However, further guarantees need Natural Language Processing (NLP) tools in addition to pattern recognition [58], that are language dependent, i.e. [10] builds the list of forbidden words in French following an incremental process without prior dictionary. Similarly, the use of an open source solution for statistical identity scrubbing is demonstrated [22] with little human effort. Another approach to impede the identification of the patient is disassociation [31] obtained by partitioning the EHR in several pieces allowing reconstruction or separate processing. When considering image data, the anonymization may require even image processing to assess that no identity information is slipped into printed images [37]. On the other hand anonymous sharing and cooperative processing of clinical and signal data via web service on a multicentre study on deep brain stimulation for Parkinson's Disease has been demonstrated [44].

### 5.3 *Interfaces*

The way that information must be entered in the EHR, and how it is managed for searches and visualization are critical issues for system acceptance. Works reported in [46, 47] provide a system where the user may define the visual workspace, aggregating data from diverse sources, and sharing the resulting visualization with other users. This system allows different strategies to manage data complexity. The ability to capture data directly from the medical devices allows real improvements of the workflow, alleviating the work on clerical tasks. In a clinical follow-up of pacemakers, remote interrogation of the devices has been demonstrated by [15]. The connection to wearable devices allows the integration in the EHR of real time measurements of physiological variables [9].

### 5.4 *Social Networks*

The eclosion of social network systems with health information [28] introduces new forms to deal with EHR information from the patient point of view, in order to self-manage health issues, allowing also amateur epidemiology, and new ways to obtain consent to do data mining on health information.

### 5.5 *Machine Learning and Computer Aided Diagnosis*

Direct learning from EHR is a challenge, because of both the heterogeneous and structured nature of the data, which is inappropriate for conventional machine learning. As an example, application of restricted Boltzmann Machines [51] required ad hoc modifications (preservation of data structure and non negative weights) in order to produce meaningful results in a predictive study about suicide. After statistical anonymization of EHR data [11] forming clusters of data for recommendation generation in an emergency environment.

## 6 Conclusions and Discussion on User Acceptance

In a general view [34] the conditions for the wide acceptance of EHR are filling the purpose from patient's point of view which is the quality of health services, smoothing the process of the professional, and warranting data quality as well as privacy for the public. Some reviews find that this acceptance is still lower within the medical practitioners [21] than in other stakeholders, such as patient advocates and administrative personnel. Social structures in the health provider ecosystem are argued to

be key for spreading adoption of EHR by social contagion [17]. There are critical reviews [12] stating that documentation takes an excessive share of clinician's time (between 25 and 50 %) with little evidence of benefit to the patient because of information overload. However, improvement on information input (i.e. tablets at point of care) have been found to have a positive impact on quality of care. Despite recognized positive publication bias [55], most clinical trial literature reports are neutral.

The current acceptance of the EHR systems is still mixed. In a recent review [38] found mixed results regarding improvements of workflow, time spent filling the EHR versus patient attendance, while there was a general acknowledgement of improving quality of information and administrative processes. A study in primary care clinics [24] finds positive correlation between EHR use and a priori attitude towards it. The main users of EHR systems are nurses, which, however, have little saying in design or purchasing decisions [20]. In emergency departments, the benefits of the EHR are more immediate than in other areas of health care, as the personnel must respond quickly to a situation with no background information of the patient. The EHR plays a central role to avoid treatment errors in situations of great stress. [2]. Training the practitioner becomes critical both for system acceptance and to achieving the expected cost/benefit results. It can be done on simulations in time critical services, such as ED [3]. The attempts to introduce cloud based PHR products by major stakeholders seem to be unsuccessful [49], because of (a) public lack of trust in the companies to store personal health information in their servers, and (b) perceived usefulness, despite promises of patient empowerment.

## References

1. Andrews, L., Gajanayake, R., Sahama, T.: The australian general public's perceptions of having a personally controlled electronic health record (pcehr). *Int. J. Med. Inform.* **83**(12), 889–900 (2014)
2. Ben-Assuli, O.: Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments. *Health Policy* **119**(3), 287–297 (2015)
3. Ben-Assuli, O., Sagi, D., Leshno, M., Ironi, A., Ziv, A.: Improving diagnostic accuracy using EHR in emergency departments: a simulation-based study. *J. Biomed. Inform.* (2015)
4. Berger, M.L., Lipset, C., Gutteridge, A., Axelsen, K., Subedi, P., Madigan, D.: Optimizing the leveraging of real-world data to improve the development and use of medicines. *Value Health* **18**(1), 127–130 (2015)
5. Bonacina, S., Marceglia, S., Bertoldi, M., Pinciroli, F.: Modelling, designing, and implementing a family-based health record prototype. *Comput. Biol. Med.* **40**(6), 580–590 (2010)
6. Burns, A.J., Johnson, M.E.: Securing health information. *IT Prof.* **17**(1), 23–29 (2015)
7. Byrne, C.M., Mercincavage, L.M., Bouhaddou, O., Bennett, J.R., Pan, E.C., Botts, N.E., Olinger, L.M., Hunolt, E., Banty, K.H., Cromwell, T.: The department of veterans affairs' (va) implementation of the virtual lifetime electronic record (vler): Findings and lessons learned from health information exchange at 12 sites. *Int. J. Med. Inform.* **83**(8), 537–547 (2014)
8. Caine, K., Lesk, M.: Security and privacy in health it [guest editors' introduction]. *IEEE Secur. Priv.* **11**(6), 10–11 (2013)
9. Castillejo, P., Martinez, J.-F., Rodriguez-Molina, J., Cuerva, A.: Integration of wearable devices in a wireless sensor network for an e-health application. *IEEE Wireless Commun.* **20**(4), 38–49 (2013)

10. Chazard, E., Mouret, C., Ficheur, G., Schaffar, A., Beuscart, J.-B., Beuscart, R.: Proposal and evaluation of fasdim, a fast and simple de-identification method for unstructured free-text clinical records. *Int. J. Med. Inform.* **83**(4), 303–312 (2014)
11. Chignell, M., Rouzbahman, M., Kealey, R., Samavi, R., Yu, E., Sieminski, T.: Nonconfidential patient types in emergency clinical decision support. *IEEE Secur. Priv.* **11**(6), 12–18 (2013)
12. Lynch, N., Kellett, J.: Medical documentation: part of the solution, or part of the problem? a narrative review of the literature on the time spent on and value of medical documentation. *Int. J. Med. Inform.* **84**(4), 221–228 (2015)
13. Danwei, C., Linling, C., Xiaowei, F., Liwen, H., Su, P., Ruoxiang, H.: Securing patient-centric personal health records sharing system in cloud computing. *Commun. China* **11**(13), 121–127 (2014) Supplement
14. De Moor, G., Sundgren, M., Kalra, D., Schmidt, A., Dugas, M., Claerhout, B., Karakoyun, T., Ohmann, C., Lastic, P.-Y., Ammour, N., Kush, R., Dupont, D., Cuggia, M., Daniel, C., Thienpont, G., Coorevits, P.: Using electronic health records for clinical research: the case of the EHR4CR project. *J. Biomed. Inform.* **53**, 162–173 (2015)
15. Diederich, L., Johnson, T.: Integrating remote follow-up into electronic health records workflow. *Health Policy Technol.* **3**(2), 126–131 (2014)
16. Fernandez-Aleman, J.L., Senor, I.C., Lozoya, P.A.O., Toval, A.: Security and privacy in electronic health records: a systematic literature review. *J. Biomed. Inform.* **46**(3), 541–562 (2013)
17. Gan, Q.: Is the adoption of electronic health record system “contagious”? *Health Policy Technol.* (2015)
18. Geissbuhler, A., Safran, C., Buchan, I., Bellazzi, R., Labkoff, S., Eilenberg, K., Leese, A., Richardson, C., Mantas, J., Murray, P., De Moor, G.: Trustworthy reuse of health data: a transnational perspective. *Int. J. Med. Inform.* **82**(1), 1–9 (2013)
19. Gkoulalas-Divanis, A., Loukides, G., Sun, J.: Publishing data from electronic health records while preserving privacy: a survey of algorithms. *J. Biomed. Inform.* **50**(0):4–19 (2014). Special Issue on Informatics Methods in Medical Privacy
20. Gleason, R.P., Farish-Hunt, H.: How to choose or change an electronic health record system. *J. Nurse Practit.* **10**(10):835–839 (2014). Special Issue: Technology That Transforms Health Care Practice and Education
21. Haluza, D., Jungwirth, D.: ICT and the future of health care: aspects of health promotion. *Int. J. Med. Inform.* **84**(1), 48–57 (2015)
22. Hanauer, D., Aberdeen, J., Bayer, S., Wellner, B., Clark, C., Zheng, K., Hirschman, L.: Bootstrapping a de-identification system for narrative patient records: cost-performance tradeoffs. *Int. J. Med. Inform.* **82**(9), 821–831 (2013)
23. Huang, L.-C., Chu, H.-C., Lien, C.-Y., Hsiao, C.-H., Kao, Tsair: Privacy preservation and information security protection for patients’ portable electronic health records. *Comput. Biol. Med.* **39**(9), 743–750 (2009)
24. Iqbal, U., Ho, C.-H., Li, Y.C.J., Nguyen, P.A., Jian, W.S., Wen, H.C.: The relationship between usage intention and adoption of electronic health records at primary care clinics. *Comput. Methods Programs Biomed.* **112**(3), 731–737 (2013)
25. Kerai, P., Wood, P., Martin, M.: A pilot study on the views of elderly regional australians of personally controlled electronic health records. *Int. J. Med. Inform.* **83**(3), 201–209 (2014)
26. Kim, D., Labkoff, S., Holliday, S.H.: Opportunities for electronic health record data to support business functions in the pharmaceutical industry—a case study from pfizer, inc. *J. Am. Med. Inform. Assoc.* **15**(5), 581–584 (2008)
27. Kukafka, R., Ancker, J.S., Chan, C., Chelico, J., Khan, S., Mortoti, S., Natarajan, K., Presley, K., Stephens, K.: Redesigning electronic health record systems to support public health. *J. Biomed. Inform.* **40**(4), 398–409 (2007). Public Health Informatics
28. Lesk, M.: Electronic medical records: confidentiality, care, and epidemiology. *IEEE Secur. Priv.* **11**(6), 19–24 (2013)
29. Li, J.: Ensuring privacy in a personal health record system. *Computer* **48**(2), 24–31 (2015)

30. Liu, J., Huang, X., Liu, J.K.: Secure sharing of personal health records in cloud computing: ciphertext-policy attribute-based signcryption. *Future Gener. Comput. Syst.* **52**, 67–76 (2014)
31. Loukides, G., Liagouris, J., Gkoulalas-Divanis, A., Terrovitis, M.: Disassociation for electronic health record privacy. *J. Biomed. Inform.* **50**(0), 46–61 (2014). Special Issue on Informatics Methods in Medical Privacy
32. Magni, A., de Oliveira Albuquerque, R., de Sousa Jr, R.T., Hans, M.G., Magni, F.G. Solving incompatibilities between electronic records for orthodontic patients. *Am. J. Orthod. Dentofac. Orthop.* **132**(1), 116–121 (2007)
33. Mertz, L.: Saving lives and money with smarter hospitals: streaming analytics, other new tech help to balance costs and benefits. *IEEE Pulse* **5**(6), 33–36 (2014)
34. Michel-Verkerke, M.B., Stegwee, R.A., Spil, T.A.M.: The six p's of electronic health records. *Health Policy Technol.* (2015)
35. Moreno-Conde, A., Jódar-Sánchez, F., Kalra, D.: Requirements for clinical information modelling tools. *Int. J. Med. Inform.* (2015)
36. Naszladay, A., Naszladay, J.: Patient health record on a smart card. *Int. J. Med. Inform.* **48**(1–3), 191–194 (1998)
37. Newhauser, W., Jones, T., Swerdloff, S., Newhauser, W., Cilia, Mark, Carver, R., Halloran, A., Zhang, R.: Anonymization of DICOM electronic medical records for radiation therapy. *Comput. Biol. Med.* **53**, 134–140 (2014)
38. Nguyen, L., Bellucci, E., Nguyen, L.T.: Electronic health records implementation: an evaluation of information system impact and contingency factors. *Int. J. Med. Inform.* **83**(11), 779–796 (2014)
39. Petrakaki, D., Klecun, E.: Hybridity as a process of technology's 'translation': customizing a national electronic patient record. *Soc. Sci. Med.* **124**, 224–231 (2015)
40. Pinciroli, F., Pagliari, C.: Understanding the evolving role of the personal health record. *Comput. Biol. Med.* **59**(0), 160–163 (2015). The State of the Personal Health Record Guest Editors
41. Puentes, J., Batrancourt, Be, Atif, J., Angelini, E., Lecornu, L., Zemirline, A., Bloch, I., Coatrieu, G., Roux, C.: Integrated multimedia electronic patient record and graph-based image information for cerebral tumors. *Comput. Biol. Med.* **38**(4), 425–437 (2008)
42. Robson, B., Caruso, T.P., Balis, U.G.J.: Suggestions for a web based universal exchange and inference language for medicine. Continuity of patient care with PCAST disaggregation. *Comput. Biol. Med.* **56**, 51–66 (2015)
43. Rosenthal, D.I.: Instant replay. *Healthcare* **1**(1–2), 52–54 (2013)
44. Rossi, E., Rosa, M., Rossi, L., Priori, A., Marceglia, S.: Webbiobank: A new platform for integrating clinical forms and shared neurosignal analyses to support multi-centre studies in parkinson's disease. *J. Biomed. Inform.* **52**(0), 92–104 (2014). Special Section: Methods in Clinical Research Informatics
45. Safadi, H., Chan, D., Dawes, M., Roper, M., Faraj, S.: Open-source health information technology: a case study of electronic medical records. *Health Policy Technol.* **4**(1), 14–28 (2015)
46. Senathirajah, Y., Bakken, S., Kaufman, D.: The clinician in the driver's seat: part 1—a drag/drop user-composable electronic health record platform. *J. Biomed. Inform.* **52**(0), 165–176 (2014). Special Section: Methods in Clinical Research Informatics
47. Senathirajah, Y., Kaufman, D., Bakken, S.: The clinician in the driver's seat: part 2—intelligent uses of space in a drag/drop user-composable electronic health record. *J. Biomed. Inform.* **52**(0), 177–188 (2014). Special Section: Methods in Clinical Research Informatics
48. Simplicio, M.A., Iwaya, L.H., Barros, B.M., Carvalho, T.C.M.B., Naslund, M.: Secourhealth: a delay-tolerant security framework for mobile health data collection. *IEEE J. Biomed. Health Inform.* **19**(2), 761–772 (2015)
49. Spil, T., Klein, R.: The personal health future. *Health Policy Technol.* **4**(2), 131–136 (2015)
50. Sujansky, W., Wilson, T.: DIRECT secure messaging as a common transport layer for reporting structured and unstructured lab results to outpatient providers. *J. Biomed. Inform.* **54**, 191–201 (2015)
51. Tran, T., Nguyen, T.D., Phung, D., Venkatesh, S.: Learning vector representation of medical objects via emr-driven nonnegative restricted boltzmann machines (enrbm). *J. Biomed. Inform.* **54**, 96–105 (2015)

52. Trcek, D., Brodnik, A.: Hard and soft security provisioning for computationally weak pervasive computing systems in e-health. *IEEE Wireless Commun.* **20**(4), 22–29 (2013)
53. Urbauer, P., Sauermann, S., Frohner, M., Forjan, M., Pohn, B., Mense, A.: Applicability of ihe/continua components for PHR systems: learning from experiences. *Comput. Biol. Med.* **59**(0), 186–193 (2015). The State of the Personal Health Record Guest Editors
54. Van Gorp, P., Comuzzi, M., Jahn, A., Kaymak, U., Middleton, B.: An open platform for personal health record apps with platform-level privacy protection. *Comput. Biol. Med.* **51**, 14–23 (2014)
55. Vawdrey, D.K., Hripesak, G.: Publication bias in clinical trials of electronic health records. *J. Biomed. Inform.* **46**(1), 139–141 (2013)
56. Vivas-Consuelo, D., Usó-Talamantes, R., Trillo-Mata, J.S., Caballer-Tarazona, M., Barrachina-Martínez, I., Buigues-Pastor, L.: Predictability of pharmaceutical spending in primary health services using clinical risk groups. *Health Policy* **116**(2–3), 188–195 (2014)
57. Zhou, J., Cao, Z., Dong, X., Lin, X., Vasilakos, A.V.: Securing m-healthcare social networks: challenges, countermeasures and future directions. *IEEE Wireless Commun.* **20**(4), 12–21 (2013)
58. Zuccon, G., Kotzur, D., Nguyen, A., Bergheim, A.: De-identification of health records using anonym: Effectiveness and robustness across datasets. *Artif. Intell. Med.* **61**(3), 145–151 (2014). Text Mining and Information Analysis of Health Documents

# Laparoscopic Video and Ultrasounds Image Processing in Minimally Invasive Pancreatic Surgeries

P. Sánchez-González, I. Oropesa, B. Rodríguez-Vila, M. Viana,  
A. Fernández-Peña, T. Arroyo, J.A. Sánchez-Margallo, J.L. Moyano,  
F.M. Sánchez-Margallo and E.J. Gómez

**Abstract** Due to limitations in conventional medical imaging and the restrictions imposed by both the anatomy and the surgical approach in pancreatic cancer, there is a need for methods to support intraoperative imaging in order to improve their accurate anatomical localization and the characterization of their nature. Laparoscopic ultrasounds (LUS) images and endoscopic videos can be used to extract useful information during the surgical procedures. A fast approach for acquiring an estimation of the tumor positioning and size through laparoscopic ultrasounds images has been developed. Based on the surgical video, endoscope 3D tracking is achieved by means of a Shape-from-Motion technique. Intraoperative imaging algorithms' validation has been carried out in an ex vivo porcine model and results have shown the viability of exploiting them for structures characterization and their 3D reconstruction.

**Keywords** Image and video guided surgery • Intraoperative imaging • Laparoscopic ultrasounds • Endoscopic video • Segmentation • 3D reconstruction • Tracking

---

P. Sánchez-González (✉) · I. Oropesa · B. Rodríguez-Vila · M. Viana ·  
A. Fernández-Peña · T. Arroyo · E.J. Gómez  
Biomedical Engineering and Telemedicine Centre, ETSI Telecomunicación,  
Universidad Politécnica de Madrid, Madrid, Spain  
e-mail: psanchez@gbt.tfo.upm.es

P. Sánchez-González · I. Oropesa · B. Rodríguez-Vila · E.J. Gómez  
Biomedical Research Networking Center in Bioengineering, Biomaterials  
and Nanomedicine, Madrid, Spain

J.A. Sánchez-Margallo · J.L. Moyano · F.M. Sánchez-Margallo  
Minimally Invasive Surgery Centre Jesus Usón, Cáceres, Spain

## 1 Introduction

Pancreatic adenocarcinoma presents the worst survival rates amongst cancers of the gastrointestinal tract. Most patients die within the first year after diagnosis [1]. Surgery is the only effective treatment, with a five-year survival rate of 20 % [2]. High mortality values are directly related to late diagnosis, since only 15–20 % cases are eligible for surgery at that stage.

Pancreatic tumor resections imply an important challenge for clinicians and oncologists. Irrespective of the approach used, surgeons face the same problems: (i) Poor correlation between the preoperative anatomical information and the actual surgical field; (ii) Incomplete information on the exact nature of the tumor; (iii) Difficulty to clearly define resectability; (iv) No guidance to perform the resection.

NAVISurg research project aims to develop an intelligent surgical navigator for pancreatic interventions [3]. The navigator is based on the Image and Video Guided Surgery (IVGS) paradigm [4], which contemplates the exploitation of the information from the endoscope to update the preoperative information of the patient to the reality of the Operating Room (OR). The navigator will additionally incorporate two clinical decision support systems (CDSS): the first one to support the diagnosis and to provide counsel on the best therapeutic approach to follow (based on patient record, preoperative studies and explorative image form LUS and video) and the second to assist surgeons for intraoperative decisions and prognosis prediction (based on endoscope video, LUS and clinical knowledge).

In order to achieve its goals NAVISurg proposes an extensive research on image processing techniques related both to the use of LUS and the endoscopic video. Both image modalities are able to provide real time intraoperative information, more so in the case of the endoscope, which is an ever-present system during the intervention. Other image modalities, such as C-arm based CTs or MR systems are commercially available. However, these systems are costly, contribute to cluttering in the OR and, since they generally need to be set up, fired and removed, may interrupt the surgical workflow. Moreover, intraoperative MR systems require demagnetized instrumental in the OR, and while modern intraoperative CT systems can keep radiation levels at a minimum, there is still the unnecessary risk of overexposure both for surgeons and patients.

This research is focused on the intraoperative imaging processing results. First, an estimation of the tumor positioning and size through LUS images will be presented. Second, based on the surgical video, endoscope 3D tracking will be achieved by means of SIFT/SURF characteristics detection techniques.

Despite the advantages of ultrasounds, segmentation of these images with or without user interaction has a high complexity because these images present a high level of speckle noise, attenuation, shadows and signal dropout [5]. In general, these problems involve over-segmentation, but there are not many previous studies that describe methods for segmentation of the pancreas structures using LUS. These algorithms are based on the level set method in combination with techniques of noise

filter [6, 7], with Gaussian filter or multiple-model particle filtering algorithm [8]. However, the literature in this area is limited and the segmentation error of the methods described is too high for the accuracy required in the pancreatic cancer surgery. Therefore, both preoperative and intraoperative characterization of the organ and the tumor present interesting scientific challenges still unsolved.

Reconstruction of the surgical scene may also be achieved from the application of computer vision techniques to the endoscopic video feed. In this case, unlike with the use of LUS, the surgical workflow needs not be interrupted. On the other hand, the information provided is limited to a superficial update of the intraoperative scenario; however, combining both LUS and video may provide valuable data for navigation and decision support within the OR.

A key step for reconstruction is the determination of the endoscope pose and movement within the surgical scenario. Shape-from-X techniques allow this without the need of modifying the endoscope with sensors [9].

Shape-from-shading employs the light intensity reflected by a physical object dependent on the geometrical relation between its surface and the source of light. These techniques exploit this physical property of the behavior of light in the surgical scene [10]. Shape-from-stereo vision refers to the ability to acquire information on the 3D structure and the relative distances on the scene from two or more intensity images taken from different viewpoints [11, 12]. However, they are not always feasible in Minimally Invasive Surgical (MIS) applications because stereoscopic endoscopes are not generally available in hospitals. Shape-from-motion techniques obtain depth information through movement analysis [13–15]. These algorithms usually compute local frame-to-frame motion feature matching and are refined in a global optimization moving backward and forward through the whole sequence (called bundle adjustment). State-of-the-art technology solves this problem in a scenario with static objects or rigid movements. The challenge lies in doing this in semi-rigid, deformable scenarios such as the proposed pancreatic resection.

## 2 Materials and Methods

### 2.1 Experimental Setup

An ex vivo porcine pancreas has been used in this study provided by the Minimally Invasive Surgery Centre Jesús Usón. The pancreas was placed stretched and fixed to a rigid surface so as to prevent shifts in its position during all imaging acquisition. For the registration of the different imaging modalities, both preoperative and intraoperative, five fiducial markers made with vitamin A capsules (Fig. 1, left), visible in both MRI and CT studies, were used. Two pseudo-tumors were simulated by injecting alginate into the pancreas head and body.

For the MRI data, we used a T2 weighted sequence with 20 coronal slices of 1 mm slice thickness with no intersection gap and  $0.67 \times 0.67$  mm spatial



**Fig. 1** (Left) Ex vivo porcine pancreas and fiducial markers inside the box trainer, including the Aurora emitter. (Right) Endoscopic video and LUS recorded by a laparoscopic surgeon

resolution using a 1.5T MR scanner (Phillips, The Netherlands). Previous to the CT study a medium contrast was injected into the pancreatic duct through the orifice of the common bile-duct and pancreatic duct. CT images (Brilliance CT, Phillips, The Netherlands) were acquired using a slice thickness of 0.1 mm and no gap between slices.

Once the MRI and CT studies were obtained, the pancreatic model was placed inside a box trainer (SIMULAP-IC05<sup>®</sup>, Minimally Invasive Surgery Centre Jesús Usón, Spain) for the intraoperative imaging. A HDI 5000 laparoscopic ultrasounds system (Phillips, The Netherlands) and endoscopic video were used as intraoperative imaging systems (Fig. 1, right). The video outputs from the ultrasounds system and the laparoscope were synchronized and recorded.

An electromagnetic tracking system (Aurora<sup>®</sup>, Northern Digital Inc., Canada) was used to record the 3D position and orientation of the sensors attached to the endoscope and LUS probe, as well as the location of the fiducial markers using a tracked pointer. LUS probe data were used for the 3D reconstruction of the intraoperative scene, endoscope data were used for evaluation of the video-based endoscope tracking, while fiducial markers' locations were used for rigid registration purposes.

## 2.2 LUS-Based 2D Segmentation

LUS images, like other ultrasounds modalities, suffer from different types of artifacts, shadowing effects and attenuation, making segmentation a demanding task.

Due to this fact, conventional methods are not enough to achieve accuracy or allow automation of the process. The study of the state of the art suggests that level set based methods are a better approach to the case [16].

The proposed solution consists on an algorithm for automatic seed searching across the image for initialization of the level set segmentation. The basis of the method is to find circular patterns in structures above possible acoustic shadows, for tumors allocated in the inside of the pancreas, or high intensity valued structures in proximity to the organ boundaries. Seeds found are stored and later used as the initial positioning for the center of active contours.

After seeds are found, the method proceeds to segment the structures using a threshold-based level set method. Despite the connectivity between the elements that shape the tumor being poor, it is possible to find an intensity relation. Threshold level sets perform best in this situation, defining the propagation of the active contour as a function of pixel intensity.

### 2.3 Free-Hand 3D Reconstruction

The objective of LUS volume reconstruction is to construct a 3D volume from a set of 2D ultrasounds frames and place it in the same coordinate space as the MRI. The volume reconstruction method uses VTK and is based on the work of Gobbi and Peters [17] and Lasso et al. [18].

The proposed method has two main stages:

- **Calculation of the affine transformation matrix** that determines the place of each frame in the volume. Aurora tracker provides the orientation and position of the electromagnetic sensor located at the tip of the LUS probe, defining the  $4 \times 4$  transformation matrices to place each LUS image in the tracker coordinate space. An additional matrix is needed to translate tracker matrices to the MRI coordinate space. This  $4 \times 4$  matrix is obtained applying a rigid registration between the markers' coordinates in Aurora space and MRI space. Final transform matrices for each LUS frame are obtained by multiplying tracker matrices by the rigid tracker-to-MRI registration matrix.
- **2D LUS frames insertion into a 3D volume** The intensity information of several pixels need to be interpolated in order to create a single 3D ultrasounds preoperative image of the surgical scene. Two methods are tested to interpolate voxels, nearest-neighbor interpolation and linear interpolation. In the first one, pixels are inserted into the nearest voxel not considering the values of neighboring points. The second one, trilinear interpolation, considers a kernel of  $2 \times 2 \times 2$  surrounding voxels and distributes pixels into them using weighted coefficients.

## 2.4 Endoscope Tracking

Based on a theoretical implementation of a Shape-from-Motion (SfM) algorithm, a tracking application able to follow laparoscope motion during surgeries is developed. The main scheme of the algorithm relies on a set of function modules that contain all the necessary processes to accomplish the following tasks:

- Calibration: estimate camera internal parameters and define intrinsic matrix ( $K$ ).
- Image preprocessing: remove noise, distortion and artifacts that degrade surgical video.
- Image characterization: detect and match distinctive features along the different scene views. For this purpose, Scale Invariant Feature Transform, SIFT, has been used to define singular image features [19].
- Motion geometry: apply SfM principles in order to determine the path described by the camera in the analyzed sequence.
- Optimization: deploy bundle adjustment procedures to compute the solution that minimizes the error value.

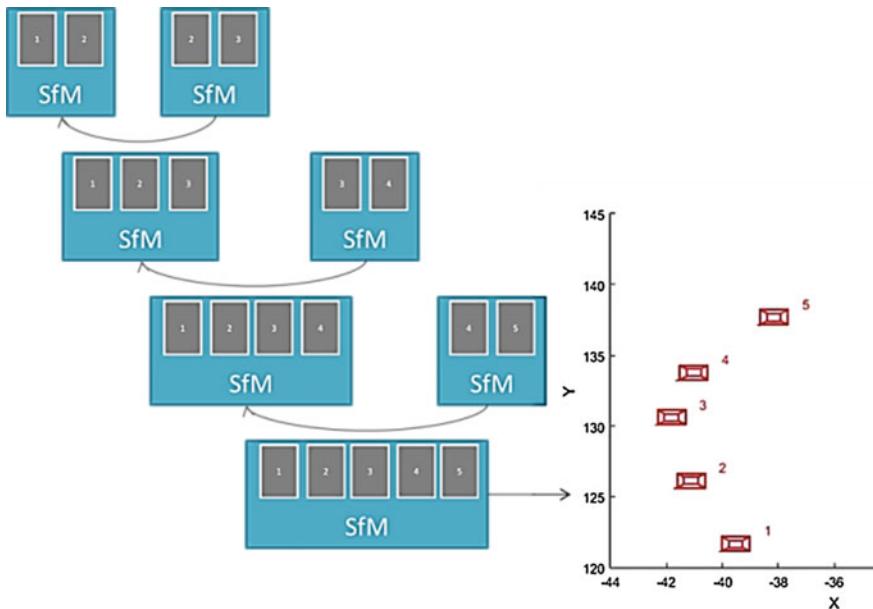
The designed model aims to solve problems related to dragging errors and to optimize the use of resources to minimize the associated computational cost. In simple terms, the model can be described as a process consisting of two levels of analysis. This analysis is sequentially applied to the current video sequence in order to obtain camera trajectory estimation.

The initial processing level pairs and analyzes the sequence's frames by keeping a common element between consecutive frame sets. This process can be interrupted if the analyzed image pair exceeds a predefined retroprojection error threshold. In this case, the conflictive pair of frames is discarded, taking up the first level process from that point in the next iteration.

The reference metric used to measure the accuracy of the estimation is the above-mentioned retroprojection error. This parameter is defined as the sum of squared errors between the measured feature points and the ones predicted by applying reverse geometry to the reconstructed points.

$$\sum_i \sum_j (\tilde{x}_i^j - K[R_i | t_i]X^j)^2 \quad (1)$$

The second level uses the estimates extracted from the pair analysis to determine the complete camera path on the sequence frame under study (Fig. 2). The initial frame defines a reference coordinate system. Camera pose estimations for the remaining views are adjusted to the reference system by successively chaining pairs to the resulting path. This is possible thanks to the shared element between consecutive pairs.



**Fig. 2** Second phase of the endoscope tracking algorithm: Pair motion estimates are combined using a common coordinate origin to compute the complete camera trajectory

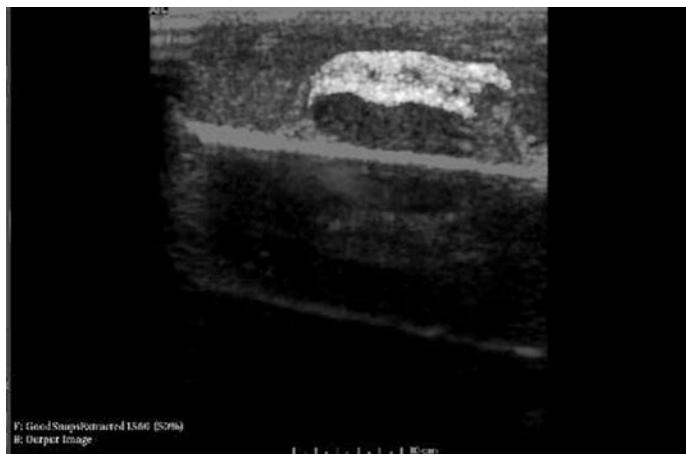
### 3 Results and Discussion

#### 3.1 LUS-Based Image Processing

The ex vivo pancreas featured two superficial pseudo-tumors. LUS video showed little to none information of these elements. Instead, shady patterns and high reflection areas were distinguishable. They also provided information of internal tissue. The exploration of the MRI of the organ confirmed presence of the two elements detected through LUS: shadows related to an air bubble between the organ and the surface and the pseudo-tumors. However, comparisons could only be made between the items not related to the latter.

Under this scenario, the segmentation algorithm was designed towards obtaining the high reflecting or shady structure caused by the pseudo-tumors and bubble alike. Defined by a range of intensities, isolation was possible. As for the volume reconstruction, areas in LUS images matching the pseudo-tumors cast big shadows underneath them. As a result, holes show up in the reconstructions giving little information of the damaged area (Fig. 3).

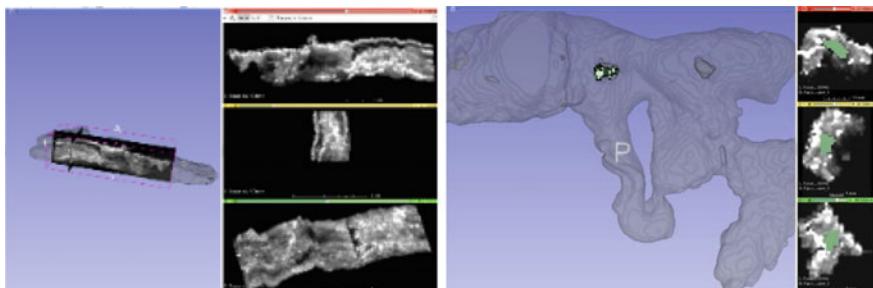
Preliminary evaluation results show that spatial correlation between the pseudo-tumors in the reconstructed LUS volume and the MRI model differ in



**Fig. 3** Recognizable structure segmentation. Highlighted area matches a shadow originated by a bubble trapped under the organ

2 mm between surfaces. This error might be caused by the pressure the probe applies to the pseudo-tumor and pancreas surface during LUS exploration.

The trilinear interpolation method took around 2 min to process a set of 1500 LUS frames. The resulting images seemed to be insusceptible to noise, but the out-coming gaps associated to the pseudo-tumors exceed their actual dimensions considerably. Nearest neighbor interpolation reconstruction method was 3 to 4 times faster. When measuring the quality of the results, the images given by the nearest neighbor interpolation contained a higher level of noise. On the other hand, it defined the pseu-do-tumors' borders more precisely. In addition, the holes the pseudo-tumor shadow originates in the volume match up quite accurately in morphology and size with the pseudo-tumor shown in the MRI (Fig. 4, left). The obtained pseudo-tumor has the same shape but differs with model in roundness in one of its sides. The obtained volume of the pseudo-tumor LUS shadow is around  $250 \text{ mm}^3$  while MRI model pseudo-tumor has a size around  $300 \text{ mm}^3$  (Fig. 4, right).

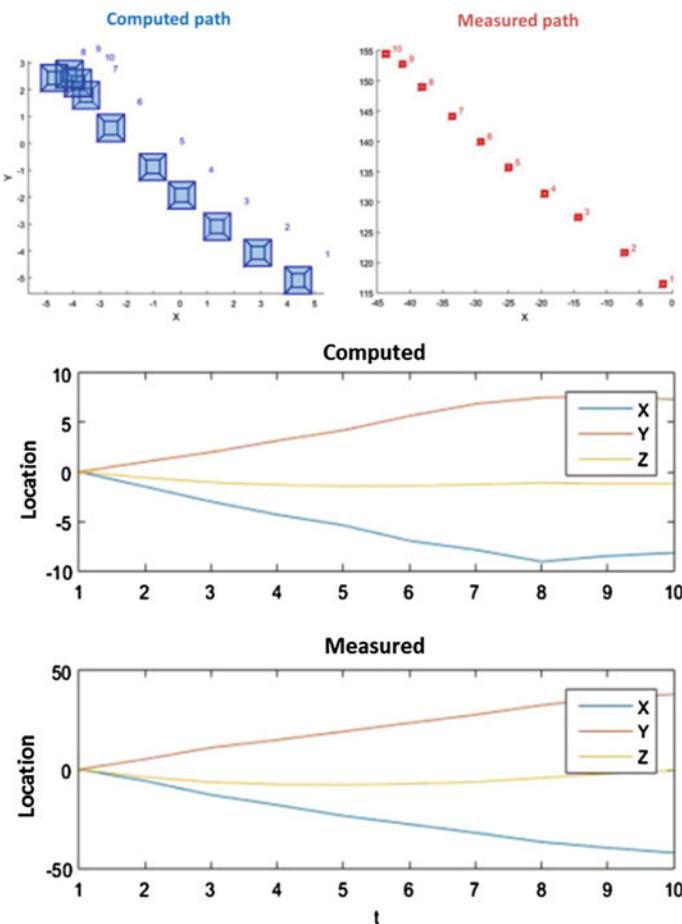


**Fig. 4** (Left) Free-hand 3D reconstruction superimposed on a pancreas model obtained from the MRI. (Right) Green areas in the reconstruction correspond to the shadow the pseudo-tumor drops

### 3.2 Video-Based Endoscope Tracking

Video sequences obtained in the experimental ex vivo setting record an environment that simulates laparoscopic interventions conditions by using different motion patterns (Fig. 1, right). Specifically, vertical, horizontal, diagonal and circular motion sequences were tested to evaluate the algorithm performance. Ground truth camera path information was obtained using the Aurora<sup>©</sup> tracking system.

To properly compare the computed camera path with the corresponding ground truth information, common temporal reference and spatial coordinate system were defined. Results show how the presented technique allows recovering laparoscope motion from surgical video analysis. Figure 5 illustrates the algorithm outcomes for a sample video footage presenting a diagonal motion pattern. Above, a comparison



**Fig. 5** Camera path reconstruction for diagonal motion video sequence. Path comparison (XY plane). 3D coordinates trajectory decomposition

of the estimated camera path and the ground truth data is represented over the XY plane. Below, the same results, computed and measured paths, are shown separately for each 3D coordinate.

The algorithm is able to follow camera motion in a qualitative manner. From a quantitative viewpoint, the estimates and reference measures are not aligned. This is due to an implicit constraint in SfM geometric principles. Dimension mismatch is not necessarily identical in all space coordinates. To correct the effect of the misalignment in the computed paths, it is necessary to apply a reverse transformation by introducing the appropriate scaling vector ( $\theta$ ). Nevertheless, it is possible to correct the scale mismatch automatically by using epipolar geometry based on the image data [20]. A semiautomatic method was introduced in the presented algorithm to reduce the negative impact in computational efficiency.

## 4 Conclusions

Results have shown the viability of exploiting intraoperative studies for structures characterization (endoscope and tumors) and their 3D reconstruction, giving surgeons new valuable and non-intrusive information during the surgical interventions. Future works will include the integration of these methods in the NAVISurg system, developed in 3D Slicer. This will simplify the visualization and comparison of images and volumes obtained from different intraoperative sources.

**Acknowledgments** This work has been carried out under project NAVISurg of the Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN).

## References

1. Beger, H.G., Rau, B., Gansauge, F., Poch, B., Link, K.H.: Treatment of pancreatic cancer: challenge of the facts. *World J. Surg.* **27**(10), 1075–1084 (2003)
2. de la Santa, L.G., Retortillo, J.A., Miguel, A.C., Klein, L.M.: Radiology of pancreatic neoplasms: an update. *World J. Gastrointest. Oncol.* **15**; **6**(9), 330–43 (2014)
3. Sánchez-González, P., Oropesa, I., Rodríguez-Vila, B., Sánchez-Margallo, J.A., Cueto, E., Peris, J.L., Mayol, J., Sánchez-Margallo, F.M., Gómez, E.J.: NAVISurg: Decision support and navigation assistance for safer image-guided surgery in pancreatic interventions. *Int. J. Comput. Assist. Radiol. Surg.* **10**(Sup.1), S226–S228 (2015)
4. Sánchez-González, P., Oropesa, I., Gómez, E.J.: Minimally invasive surgical video analysis: a powerful tool for surgical training and navigation. *Stud. Health Technol. Inform.* **190**, 33–35 (2013)
5. Michailovich, O., Tannenbaum, A.: Segmentation of medical ultrasound images using active contours in image processing. *ICIP IEEE Int. Conf. Imag. Proc.* **5**, 513–516 (2007)
6. Supriyanto, E., Mahani Hafizah, W., Wei Yun, W., Jamlos, M.: Ultrasound pancreas segmentation: a new approach towards detection of diabetes mellitus. In: Proceedings of the 15th WSEAS International Conference on Computer, pp. 184–188 (2011)

7. Mahani Hafizah, W., Wei Yun, W., Supriyanto, E.: Optimization of pancreas measurement techniques based on ultrasound images. *WSEAS Trans. Biol. Biomed.* **8**(4), 135–144 (2011)
8. Angelova, D., Mihaylova, L.: Contour segmentation in 2D ultrasound medical images with particle filtering. *Mach. Vis. Appl.* **22**(3), 551–561 (2011)
9. Maier-Hein, L., Mountney, P., Bartoli, A., Elhawary, H., Elson, D., Groch, A., Kolb, A., Rodrigues, M., Sorger, J., Speidel, S., Stoyanov, D.: Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. *Med. Image Anal.* **17**, 974–996 (2013)
10. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from Shading: a Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 690–706 (1999)
11. Samaras, D., Metaxas, D., Fua, P., Leclerc, Y.G.: Variable Albedo Surface Reconstruction from Stereo and Shape from Shading. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)* vol. 1, p. 1480 (2000)
12. Banks, J., Bennamoun, M., Kubik, K., Corke, P.: An accurate and reliable stereo matching algorithm incorporating the rank constraint. In: *Symposium on Intelligent Robotic Systems*, pp. 23–32 (1999)
13. Newman, P.M.: On the structure and solution of the simultaneous localization and map building problem. Ph.D. dissertation, University of Sydney, 1999
14. Jebara, T., Azarbayejani, A., Pentland, A.: 3D Structure from 2D Motion. *IEEE Signal Process. Mag.* **16**, 66–84 (1999)
15. Dellaert, F., Seitz, S., Thorpe, C., Thrun, S.: Structure from motion without correspondence. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)* (2000)
16. Ahmad, A., Cool, D., Chew, B.H., Pautler, S.E., Peters, T.M.: 3D segmentation of kidney tumors from freehand 2D ultrasound. In: Cleary, K.R., Galloway, Jr, R.L. (eds.) *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*, SPIE Proceedings, vol. 6141
17. Gobbi, D.G., Peters, T.M.: Interactive intra-operative 3d ultrasound reconstruction and visualization. *MICCAI* (2), 156–163 (2002)
18. Lasso, A., Heffter, T., Rankin, A., Pinter, C., Ungi, T., Fichtinger, G.: PLUS: Open-source toolkit for ultrasound-guided intervention systems. *IEEE Trans. Biomed. Eng.* **61**(10), 2527–2537 (2014). doi:[10.1109/TBME.2014.2322864](https://doi.org/10.1109/TBME.2014.2322864)
19. Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **60** (sup 2), 91–110 (2004)
20. Sturm, P., Triggs, W.: A factorization based algorithm for multi-image projective structure and motion. In: *4th European Conference on Computer Vision (ECCV'96)*, vol. 1065, pp. 709–720, Springer, Cambridge, United Kingdom

# **Erratum to: Hygehos Ontology for Electronic Health Records**

**Naiara Muro, Eider Sanchez, Manuel Graña, Eduardo Carrasco,  
Fran Manzano, Jose María Susperregi, Agustín Agirre  
and Jesús Gómez**

## **Erratum to:**

**Chapter “Hygehos Ontology for Electronic Health  
Records” in: Y.-W. Chen et al. (eds.), *Innovation  
in Medicine and Healthcare 2016*, Smart  
Innovation, Systems and Technologies 60,  
DOI [10.1007/978-3-319-39687-3\\_33](https://doi.org/10.1007/978-3-319-39687-3_33)**

In the original version of the book, in Chap. 30, Figure 3 was affected and has to be replaced. The erratum chapter and the book have been updated with the change.

---

The updated original online version of this chapters can be found at  
[http://dx.doi.org/10.1007/978-3-319-39687-3\\_30](http://dx.doi.org/10.1007/978-3-319-39687-3_30)

# Author Index

## A

- Agirre, Agustin, 311  
Alamis, Arnulfo, 93  
Alanis, Arnulfo, 99, 111  
Alberich, Maider, 291  
Alcaraz, Raúl, 3  
Alicante, Anita, 183  
Alonso, Jesús B., 25  
Amoruso, Chiara, 137  
Arlati, Sara, 147  
Arroyo, T., 333  
Artetxe, Arkaitz, 291  
Asselbergs, Joost, 37

## B

- Baltazar, Rosario, 99, 111  
Barreira, N., 279  
Blaunstein, N., 83  
Brancati, Nadia, 137  
Bumbea, Ana-Maria, 259

## C

- Calvo, C., 279  
Carmona-Duarte, Cristina, 25  
Caro, Maricela Sevilla, 123  
Carrasco, Eduardo, 311  
Casillas, Miguel A., 99, 111  
Chen, Yen-Wei, 209, 237, 227  
Chyzyk, Darya, 301  
Cohen, Y., 83  
Corazza, Anna, 183  
Corda, Daniela, 137

## D

- De Pietro, Giuseppe, 137, 269, 171  
Dekel, B., 83  
del Carmen Osuna Millán, Nora, 129  
del Consuelo Salgado Soto, María, 129

Dhillon, Baljean, 71

Dong, Jiangtao, 197

## E

- Echaniz, Oier, 323  
Esposito, Massimo, 171

## F

- Fernández-Aguilar, Luz, 15  
Fernández-Caballero, Antonio, 3, 15  
Fernández-Peña, A., 333  
Fernandez-Varela, Isaac, 61  
Ferrer, Miguel A., 25  
Fontana, Luca, 147  
Forastiere, Manolo, 269  
Franchin, Cristina, 147  
Frucci, Maria, 137  
Fu, Lanxing, 71  
Fujii, Ryoma, 209  
Fujimoto, Yuta, 237  
Funahashi, Koji, 217

## G

- García-Martínez, Beatriz, 3  
García-Tarifa, María J., 249  
Garza, Arnulfo Alanís, 123  
Gómez, E.J., 333  
Gómez, Jesús, 311  
Gómez-Vilda, Pedro, 25  
González, Felma, 93  
Górriz, Juan M., 249  
Goto, Tomio, 217  
Gramigna, Cristina, 147  
Graña, Manuel, 301, 323, 311

## H

- Han, Xian-Hua, 227  
Hasegawa, Kyoko, 237

Hermida, A., 279

Hernandez-Pereira, Elena, 61

Hoogendoorn, Mark, 37

Hu, Hongjie, 227

## I

Isgrò, Francesco, 183

Ivanescu, Mircea, 259

## K

Karacapilidis, Nikos, 159

Kariya, Hidetoshi, 217

Kitrungrotsakul, Titinunt, 209

Konno, Yu, 227

## L

Latorre, José Miguel, 15

Lemus, Lenin, 111

Li, Jingbing, 197

Lin, Lanfen, 227

Liu, Yitao, 227

Llamas, Teresa Barrón, 111

Lobato, Bogart Yail Márquez, 129

Londral, Ana Rita M., 25

López, J.E., 279

López, María T., 15

López, Sandra, 99

Lozano, Francisco, 249

## M

Maclair, Grégory, 291

Magdaleno-Palencia, Jose Sergio, 93

Malosio, Matteo, 147

Manzano, Fran, 311

Márquez, Bogart Yail, 93

Martínez-Murcia, Francisco J., 249

Martínez-Rodrigo, Arturo, 3, 15

Mastropietro, Alfonso, 147

Méndez, Guillermo, 99

Mendez-Mota, Sergio, 93

Minutolo, Aniello, 171

Molteni, Franco, 147

Moreno, Hilda Beatriz Ramírez, 123

Moret-Bonillo, Vicente, 61

Mori, Takuma, 217

Moshtael, Howard, 71

Moyano, J.L., 333

Mrakic-Sposta, Simona, 147

Mucic, Davor, 49

Muro, Naiara, 311

## N

Nonell-Canals, Alfons, 291

Novo, J., 279

## O

Oropesa, I., 333

Ortega, M., 279

Ortiz, Andrés, 249

## P

Palencia, José Sergio Magdaleno, 123, 129

Pastor, Johnno, 37

Pastor, José Manuel, 15

Penedo, M.G., 279

Petrisor, Anca, 259

Pittaccio, Simone, 147

Pla, Miguel A. Mateo, 99

Plamondon, Réjean, 25

Polishchuk, G.S., 83

Popescu, Dorin, 259

Popescu, Livia-Carmen, 259

Popescu, Razvan, 259

Potamias, George, 159

## R

Ramírez, Javier, 249

Ramírez, Margarita Ramírez, 123, 129

Ricarte, Jorge, 15

Riper, Heleen, 37

Ritter, Gerhard, 301

Rizzo, Giovanna, 147

Rodríguez-Vila, B., 333

Rojas, Esperanza Manrique, 123, 129

Romo, Karina, 93

Ros, Laura, 15

Ruwaard, Jeroen, 37

## S

Sacco, Marco, 147

Sakurai, Masaru, 217

Sanchez, Eider, 311

Sánchez-González, P., 333

Sánchez-Margallo, F.M., 333

Sánchez-Margallo, J.A., 333

Sannino, Giovanna, 269

Santamaría-Navarro, Eduardo, 291

Sergeev, M.B., 83

Serrano, Juan Pedro, 15

Shimizu, Masato, 217

Silvestri, Stefano, 183

Susperregi, Jose María, 311

## T

Tanaka, Satoshi, 209, 237

Tateyama, Tomoko, 209, 237

## U

Underwood, Ian, 71

**V**

- van Breda, Ward, [37](#)  
Varlamova, L.L., [83](#)  
Varone, Alessia, [137](#)  
Viana, M., [333](#)

**X**

- Xu, Rui, [237](#)

**Z**

- Zamudio, Víctor, [99, 111](#)  
Zangróniz, Roberto, [3, 15](#)  
Zhu, Wenchao, [227](#)  
Zilberman, A., [83](#)  
Zuñiga, Lenin Lemus, [99](#)