

Advances in Pattern Recognition

Research in Computing Science

Series Editorial Board

Editors-in-Chief:

Grigori Sidorov (Mexico)
Gerhard Ritter (USA)
Jean Serra (France)
Ulises Cortés (Spain)

Associate Editors:

Jesús Angulo (France)
Jihad El-Sana (Israel)
Alexander Gelbukh (Mexico)
Ioannis Kakadiaris (USA)
Petros Maragos (Greece)
Julian Padget (UK)
Mateo Valero (Spain)

Editorial Coordination:

Alejandra Ramos Porras

Research in Computing Science es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 147, No. 11**, noviembre de 2018. Tiraje: 500 ejemplares. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No.: 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Impreso en la Ciudad de México, en los Talleres Gráficos del IPN – Dirección de Publicaciones, Tres Guerras 27, Centro Histórico, México, D.F. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

Editor responsable: *Grigori Sidorov, RFC SIGR651028L69*

Research in Computing Science is published by the Center for Computing Research of IPN. **Volume 147, No. 11**, November 2018. Printing 500. The authors are responsible for the contents of their articles. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research. Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

Advances in Pattern Recognition

J. Arturo Olvera-López

J. Ariel Carrasco-Ochoa

J. Francisco Martínez-Trinidad

Sudeep Sarkar (eds.)



Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"



Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2018

ISSN: 1870-4069

Copyright © Instituto Politécnico Nacional 2018

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional “Adolfo López Mateos”, Zacatenco
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Printing: 500

Printed in Mexico

Editorial

This volume of the “Research in Computing Science” journal (RCS) contains selected contributions related to either Pattern Recognition principles or Pattern Recognition Applications research fields. In particular, they were presented in the 6th MCPR2018-PSM at MCPR2018 conference.

The contributions published in this volume were carefully evaluated by members of a Technical Committee who are experts in the Pattern Recognition Field and related areas.

In this volume, the topics covered by the contributions are: Bio-signal Analysis interfaces, 3D Object Reconstruction, EEG Spectrograms Analysis, Emotion Classification, and Breast Thermographic Image Segmentation.

We cordially thank to the RCS editorial board for allowing the opportunity of publishing this volume. In addition, a special acknowledgment to CONACYT (Mexico) for the support obtained through the project DADC-292930. Finally, we thank to our reviewing Committee for their valuable participation as well as to authors for their submitted contributions.

We hope the contributions in this volume will be useful to the reader interested in Pattern Recognition, their applications and related areas.

J. Arturo Olvera-López
J. Ariel Carrasco-Ochoa
J. Francisco Martínez-Trinidad
Sudeep Sarkar

Guest Editors

2018

Table of Contents

	Page
Towards a Method for Biosignals Analysis as Support for the Design of Adaptive User-Interfaces	9
<i>Heber Avalos-Viveros, Guillermo Molero-Castillo, Edgard Benitez-Guerrero, Everardo Bárcenas</i>	
Phase Unwrapping for 3D Object Reconstruction by means of Population-based Metaheuristics	21
<i>Daniel Duarte-Carrera, Alfonso Rojas-Domínguez, Luis Carlos Padierna</i>	
Time-Frequency Analysis of EEG Spectrograms using 2-D Gabor Filters for Epileptic Seizure Classification	31
<i>Ricardo Ramos-Aguilar, J. Arturo Olvera-López, Ivan Olmos-Pineda, Manuel Martín-Ortiz</i>	
Emotion Classification of Twitter Data Using an Approach Based on Ranking	45
<i>Cecilia Reyes-Peña, David Pinto-Avendaño, Darnes Vilariño-Ayala</i>	
Automatic Segmentation in Breast Thermographic Images Based on Local Pattern Variations	53
<i>Daniel Sánchez-Ruiz, Iván Olmos-Pineda, J. Arturo Olvera-López</i>	

Towards a Method for Biosignals Analysis as Support for the Design of Adaptive User-Interfaces

Heber Avalos-Viveros¹, Guillermo Molero-Castillo^{1,2},
Edgard Benitez-Guerrero¹, Everardo Bárcenas^{1,2}

¹Universidad Veracruzana, Xalapa, Veracruz, Mexico
heeber.av@gmail.com, edbenitez@uv.mx

²CONACYT–Universidad Veracruzana, Mexico
{ggmolero, iebarcenas}@conacyt.mx

Abstract. Biosignals are information sources obtained from the different biological and physiological structures of the human organism. This paper presents the advance of the proposal of a method for the biosignals analysis, which allows serving as support for the design of adaptive user-interfaces. This method includes four stages: biosignals collection, extraction and preparation, analysis, and getting patterns. This analysis and obtaining of user patterns through the biosignals could be especially useful because they represent valuable information related to events or actions of user behavior, which could be incorporated in the stage of requirements specification for the design of adaptive user-interfaces, and even serve for the refinement of these.

Keywords: Biosignals, Adaptive user-interfaces, Data mining, User requirements.

1 Introduction

In recent years, the technological development has allowed obtaining data sources through different devices, as brainwave diadems, smartwatches, smartphones and even by other body sensors. This in order to obtain biosignals generated by the human body [24]. These devices have processing capabilities, which make it possible the continuous collection of data flows about users and their environment, as interactions, locations, physiological states, contact with other users and other digital traces [27].

In this context, biosignals could be collected into different frequencies, with or without intervention human, and for prolonged periods [28]. Nevertheless, it requires a series of analytical considerations, beginning with careful experimentation to ensure the analysis and design of a case study, participant training, and data storage [27].

Therefore, through these biosignals, we try to know the human behavior related to events, that is, through the analysis of data we could discover certain events inherent in the user and their environment [8]. Through this type of

analysis, data patterns related to the user could be obtained through the application of data mining techniques, as clustering, association rules, decision trees, Bayesian networks, among others.

For example, a case of special interest that is addressed in this research is to find a way to capture, process and analyze biosignals of the user in order to obtain information of interest, known as patterns, for the design of adaptive interfaces. These user data patterns can serve as support information for the design of adaptive user-interfaces of a particular system.

However, to have truly adaptive interfaces is necessary to acquire diverse information about the user, as your preferences, needs, features and software interests [14]. This represents a research challenge in human-computer interaction (HCI) for the detection and analysis of biosignals, reliably, related to human behavior events [15].

One of the current approaches to the capture of biosignals is through low-cost devices, as a) Emotiv Epoc, for the capture of brainwave signals through a multichannel electroencephalogram (EEG), b) smartwatches, to measure heart rates, and c) smartphones, for measure and control cardiac pulses through optical sensors, like the electrocardiogram. At present, these devices have a wide range of applications in HCI and interfaces customization [15]. Therefore, as part of this research, we have the intention to collect biosignals through this type of devices, which could be a complex task due to the noise of the signal [3].

This paper presents an advance of the proposal of a method for the biosignals analysis as support the design of adaptive user-interfaces. This method is divided into four phases: a) collection, b) extraction and preparation, c) analysis, d) getting patterns. It is sought that the information obtained from the user, through the biosignals analysis, could be useful as a support for the definition of requirements in the design of the adaptive interfaces. Since this is usually a complex task because it is needed translate the software needs of the users into a functions set and restrictions.

2 Background

2.1 Biosignals

A signal is a form of data transmission, the acquisition of which allows obtaining information about the source that generated it. In this sense, acquisition sources of biosignals are the different biological and physiological structures of the human body organism [13]. The representation of the biosignals facilitates the analysis and identification of the data. The main aim of the processing and analysis of these signals is the information acquisition, diagnosis, monitoring, therapy, control, and evaluation [22].

The capture of these biosignals allows us to extract information about the functioning of the different organs [13]. In the biosignals, amplitude and frequency ranges are considered as significant factors [10]. These ranges may vary depending on the method of acquisition, whether through brainwave diadems,

smartwatches, smartphones, among other devices, and their values are approximate, whether they are normal or out of range. Figure 1 shows a conceptual representation of the extraction and analysis of biosignals in order to serve as support information in the definition of requirements for the design of adaptive user-interfaces.

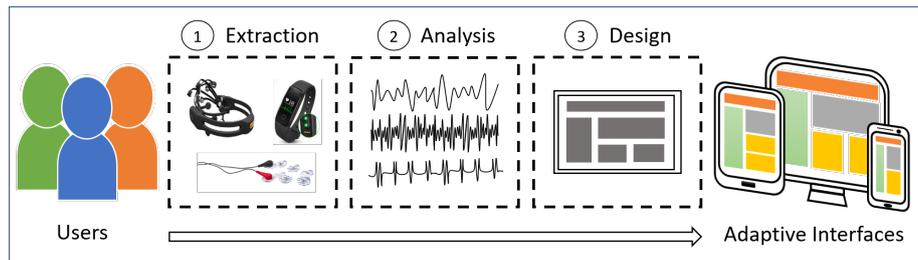


Fig. 1. Conceptual representation of acquisition and analysis of biosignals.

Given the variety of biosignals, nowadays, there are at least three ways to classify them [13]: by their existence, nature, and origin. By its existence: a) permanent, which exist without any artificial impact, or excitement; b) induced, which are triggered artificially, excited or induced. By its nature: a) quasi-static, that transports information in its stationary state that can exhibit relatively slow changes with time; b) dynamics, which produce large changes in the time domain. By its origin: a) electrical, such as the nervous system and muscle cells; b) magnetic, such as the heart, the brain, the lungs, among others; c) mechanical, such as movement, displacement, tension, force, pressure and flow; d) optic, which is generated from the optical attributes of biological systems; e) acoustic, blood flow, muscle noises; and f) chemical, which contain information about the levels and changes of chemical agents in the body.

2.2 Interfaces

Before 1960, the term user interfaces was practically non-existent. Later, in 1963 Ivan Sutherland published, in his PhD thesis, the creation of graphic images directly on a screen using an optical pencil. It was the appearance of the first graphical user interface [17]. Then, in the early 1960s, Douglas Engelbart and William English worked on a project on a new control device, which was the first mouse. Also, in 1968 Engelbart presented a system that was the first to incorporate hypertext, windows, document control and teleconferences.

On the other hand, in the 80s, the evolution of HCI became a field of user-centered research [12]. This field of knowledge studies the design, evaluation, and implementation of interactive technological devices, where technology happened to become a support tool [2].

Nowadays, user interfaces are the part of the system that is seen, heard and felt. They are the means by which the user can communicate with a device or computer. This way, HCI takes place through the user interface, which consists of interaction through a screen, keyboard, mouse, speaker, special buttons, lights, gloves that can feel the movements of the fingers, eye sensors, among others [11].

In this sense, adaptive interfaces are those that adapt automatically to the characteristics of the users, allowing the improvement in the satisfaction and permanence of the user interacting with the system [21]. Current trends are towards systems for the recovery and refinement of information [16], this depending on the interaction patterns that allow adapting the results through adaptive interfaces. However, at present it is sought interfaces that may be understandable, seeking greater effectiveness when performing certain tasks [21]. For this, we seek to define new support methods for the development of adaptive interfaces in terms of user-centered design.

2.3 Current Challenges

Nowadays, the biosignals analysis has an increase in the different fields and applications. The interest consists in analyzing physiological measurements [23]. Among the limitations are the problems of handling large amounts of data, noise and other factors that affect cognitive states [26].

This situation leads to various research challenges that are classified as technical and usability [1]. The technical challenges are related to the obstacles of the system, especially in the obtaining of signals, since the data obtained are not shown linearly, but dynamically with varied ranges and frequencies and with noise in the signals. Usability challenges describe the problems that affect the level of acceptance and express limitations that users face when using technology.

In the case of biosignals processing, it requires at least three states [4]: a) measurement (biosignals acquisition), b) transformation (redundancy elimination and noise), and c) interpretation (logical base, heuristic reasoning, and statistical origin). To realize these stages requires effort, not only by the biosignals analyst but also by the users, since the changes in mood can trigger unreliable data records.

Another challenge is the selection of an adaptive channel for users, and thus have appropriate interfaces according to the needs for each of these [14]. Here, the environment has an important role when signals are captured to avoid noise and high ranges [3]. Unfortunately, there are noise problems at the time of signal acquisition.

On the other hand, in the context of user interfaces, the search for new interfaces is associated with technological advances. This is because of every day the number of users with different training, social level and abilities increases [19]. Therefore, there is a greater willingness to use new technologies. This situation represents a challenge for the HCI, which must keep pace with these advances and make the most of the technologies to be made available to users [2].

In this sense, the construction of user-interfaces that are easy to use, easy to memorize, efficient in terms of use and with a low error rate represents another

important challenge [6]. Therefore, it is important to involve the user from the early stages of the construction process, considering aspects of interest, as the tasks, its interests, preferences, and even its physical limitations.

2.4 Related Work

At present, there are several works that seek to provide new mechanisms and applications that allow analyzing the behavior of users in certain contexts and thus develop new tools and interfaces in HCI.

In [7] the authors looked for a way to design adaptive interfaces through the biosignals analysis obtained from the human body. For this, they used functional near-infrared spectroscopy (FNIRS). This technology allowed recording changes in blood oxygenation in the brain is a non-invasive way and to obtain measurements in environments with lots of light and noise. By analyzing some interfaces, it was possible to measure the workload of the users when performing certain tasks. Two types of users were used, those who had knowledge about the operation of the interface and those who did not have prior information. The results determined that 90% of users with knowledge about the interface presented less workload compared to those who did not have knowledge. With this, the authors determined that the functioning of an interface could be measured through FNIRS.

Another approach to the design of adaptive user-interfaces was proposed by [15], where they developed a framework that allows acquiring related information about the interests and profiles of users. They place special emphasis on the design of adaptive user-interfaces for people with some type of limitation, that is, due to cerebrovascular accidents or aging, in order to improve accessibility problems in this type of users.

In [9] the biosignals collection was sought, but in a multimodal context, that is, acoustic signals (ACC), electromyography (EMG) and electroencephalogram (EGG) were used. The authors designed a framework for obtaining of signals through various devices at the same time. The objective was to create a tool for the collection, storage, and visualization of the signals obtained from the devices, this in order to analyze them based on possible relationships and dependencies between them. To later be used in the development of adaptive interfaces.

In [3] they used electroencephalography (EEG) signals from a point of view of experimental psychology and cognitive science. They focused on this type of signals because there are currently low-cost devices compared to traditional devices used in the field of Health. For this, they used Emotiv EPOC device, which allows obtaining information through brain waves. Two studies were done, one to determine the feasibility of using Emotiv EPOC to measure interhemispheric spectral differences, and another to analyze variations in cerebral activity. The results showed that, for the first case, the data obtained were sufficiently clean, which allowed revealing characteristics of interest through the EEG, while in the second case the variations in the signals were demonstrated by physical stimuli of the user.

In [20] the possibilities of using biosignals processing in real time were explored. This to improve the interaction with the user-interfaces. The analysis of an interface was made through an intelligent agent, who oversaw asking the user questions in order to know their reactions through eye signals and brain waves. The results obtained allowed to know the physical reactions generated by the user, this allowed to reflect and adapt their behavior according to the user's emotions.

In [5], the evaluation of adaptive user-interfaces was sought, taking into account the emotional state and workload of the user when interacting with these. For this, monitoring devices were used that allowed to capture physiological signals of the human body, such as eye-tracking and brain waves through Emotiv EPOC. The study focused on measuring the time it took the user to use an interface with certain tasks, as well as measuring the number of errors that arose at the time of their interaction. These tasks were carried out through three interaction mechanisms: a) toolbar, b) menus, and c) interactive bar, named Boulevard. The results determined that the user had a good acceptance and little workload when using the proposed interaction style (Boulevard). Therefore, they demonstrated that the use of physiological signals measuring devices could be of great help to know the behavior of the user with certain user interfaces.

Finally, in [29], the authors sought to provide a graphical user-interface as an expressive channel to represent the artistic qualities of several users with a disability. They presented special attention in the design of the interface, whose input data were biosignals type electromyography (EMG), electrooculography (EOG) and electroencephalography (EEG). For this, they used a mechanism through which the biosignals were acquired, then the data processing was done, where to mitigate low signals and with noise, was used the mathematical method called Lorenz System. Thus, through the results obtained from various tests with users, with and without disabilities, they were able to determine that, through biosignals as data entry in the interfaces, it was possible to execute actions in the application; in this case they used predefined data to model a vase, and then design one freely.

3 Method

In order to obtain reliable biosignals data sources, as support for the design of adaptive interfaces, four work stages were defined, shown in Figure 2. These stages are exploratory and applied type since it is an emerging issue that seeks to solve the problem through a theoretical-practical exploration.

In the first stage, the beginning of the research is proposed, which is fundamental to obtain reliable data sources that support the analysis of biosignals obtained from physical and logical sensors, with the purpose of measuring physiological signals of the users. The second stage focuses on the analysis and conceptual design of the solution proposal for obtaining and analyzing biosignals. The third stage is the development of the analysis method associated with the study case. The fourth stage focuses on the evaluation of the biosignals analysis

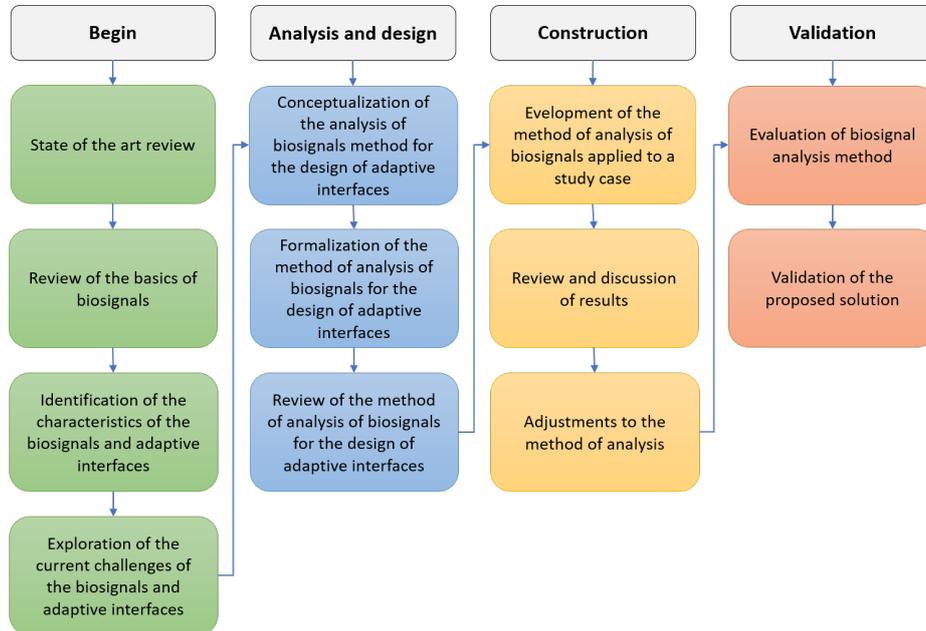


Fig. 2. Method of work for the development of the research proposal.

method, from the point of view of performance and compliance with the initial requirements and objectives.

4 Proposal

The analysis and design of user-interfaces is a complex process since the requirements and needs of the users must be known, to whom the interface is addressed [7, 15]. Currently, to identify these requirements and needs of the users, conventional techniques are used [18], as documentation analysis, observations, interviews, quizzes, mind maps, brainstorming, sketches, prototypes and other instruments of interest.

However, the previous techniques may not be enough, due to poorly used the development tools and methods by software developers [25]. Therefore, an adequate understanding of the user's requirements is of great importance, in this way, it is fundamental to consider other aspects of the users, such as: their interests, preferences and even affective states, as well as the development environment and the available technologies.

Based on the above, this research project seeks to design a method to extract and analyze biosignals data obtained from users, with the purpose of using this information as support in the design of adaptive user-interfaces. The reason for obtaining and analyzing biosignals is that they represent valuable information related to events or actions of user behavior. For this, as a method four stages

of work are proposed (see Figure 3) a) biosignals collection, b) extraction and preparation, c) analysis, and d) getting patterns.

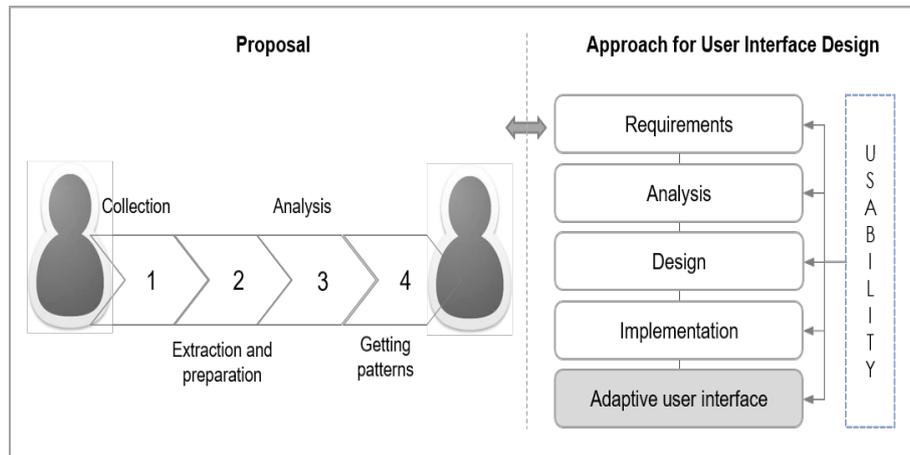


Fig. 3. Conceptual design of the biosignals analysis method.

The collection (phase 1) involves the use of devices to capture the EEG type biosignals and cardiac pulses, either through brainwave tunes, smartwatches and even by body sensors. Once the acquisition devices have been defined, it must be defined a data capture strategy, for this, it is suggested to deliver to the users a letter of consent and a fact sheet with the activities that must perform on a specific user-interface. For the execution of activities, it is recommended to use the guided tasks method, that is, the participants will be dictated aloud to tasks they must perform. As part of the strategy, it is also suggested recording the tests in order to monitor the interaction of users with the application. This with the purpose of finding usability problems in the user interface.

In the extraction and preparation (phase 2) a selection of data is made to obtain a base source of biosignals on which subsequent analyzes will be made. This selection and preparation are made because a large amount of data is commonly collected, whose measurements contain noise and atypical values, which need to be filtered, deleted and even transformed for the application of a particular technique. In addition, this type of biosignals are registered in small timestamps as seconds and milliseconds, and even in nanoseconds. Subsequently, to eliminate redundancies and possible high dimensionalities it is advisable to do an analysis of correlations and dispersion of the data.

Biosignals analysis (phase 3), apart from statistical measurements as mean, standard deviation, variance, covariance and other interest estimates, seeks to identify patterns in the data, which were previously extracted and prepared. This patterns identification is done through the use of one or more data mining techniques, as clustering (hierarchical and partitional), association rules, corre-

lations, regressions, decision trees, and artificial neural networks. These patterns represent a set of interest characteristics about the users evaluated, as tastes certain colors, sizes, format styles, presentation and other aspects of interest.

Subsequently, through the getting of patterns (phase 4), it is sought to interpret the patterns of users previously identified. The preferences of the users about a certain type of interface are described, as colors, text size, size of the icons, the position of the elements and other characteristics of interest. This information from users could be especially useful as part of the definition of requirements in the analysis and design of adaptive interfaces. Since the definition of these requirements is usually a complex task because it is necessary to translate the software needs of users into a set of functions and restrictions.

In this sense, considering the approach of developing adaptive user-interfaces (requirements specification, analysis, design and implementation), the analysis of biosignals is intended to support the first stages of development, such as requirements specification, analysis, and design. In addition, these could be refined by the biosignals. Consequently, biosignals could be useful not only for the development team but also for the user for whom the interface is intended. Therefore, it is sought through the biosignals analysis to serve as a support mechanism for the design of adaptive user-interfaces.

5 Conclusions

Technological advances in recent years have allowed the implementation of new mechanisms for the design of adaptive user-interfaces, which seek to cover a greater degree of satisfaction of users when interacting with them. Currently, various techniques are used for the analysis and definition of requirements, which allow the construction of prototypes that could be evaluated and determine changes or improvements to the system, which is a practical method that results in an executable software model.

Another practical way to include valuable information in the specification of requirements is through the analysis and understanding of biosignals obtained from users. For its capture, diverse devices exist that allow monitoring their behavior related to the events. For this reason, it is important to consider new mechanisms to find useful information to support the design of adaptive interfaces.

As part of the research, this document showed the advance of the design of a method for the analysis of biosignals as support for the design of adaptive user interfaces. This method consists of four stages: a) biosignals collection, b) extraction and preparation, c) analysis, d) getting patterns. This proposal could be especially useful because the information obtained from users could be incorporated into the requirements specification phase of adaptive user-interface design.

As future work of this proposal, it is contemplated to make a practical experimentation on the biosignals analysis to a group of users. The objective is to make an incremental and systematic development of the proposal of extraction

and analysis of biosignals as support for the design of adaptive user-interfaces.

Acknowledgments. This work was partially supported by the National Council of Science and Technology (CONACYT) of Mexico, as part of the project Cátedras CONACYT "Infraestructura para agilizar el desarrollo de sistemas centrados en el usuario" Ref. 3053. The first author was also supported by CONACYT through a MSc. scholarship, Ref. 629916.

References

1. Abdulkader, S.N., Atia, A., Mostafa, M.S.M.: Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal* 16(2), 223–226 (2015)
2. Acosta, A., Zambrano, N.: Importancia, problemas y soluciones en el diseño de la Interfaz de Usuario. Universidad de Oriente (2006)
3. Burnet, D.H., Turner, M.D.: Expanding EEG research into the clinic and classroom with consumer EEG Systems. Georgia state university (2017)
4. Cuenca, J.N.: Informática médica: Análisis de bioseñales (2008), <http://slideplayer.es/slide/4617122/>
5. Eichler, Z.: Towards Multimodal Adaptive User Interfaces. Ph.D. thesis, Masarykova univerzita, Fakulta informatiky (2014)
6. Fernández, M.J., Ullate, J.M.A., Salvador, J.A.: Interfaces de usuario: diseño de la visualización de la información como medio de mejorar la gestión del conocimiento y los resultados obtenidos por el usuario. In: *Actas del V Congreso ISKO-España*. pp. 42–54 (2001)
7. Girouard, A., Solovey, E.T., Hirshfield, L.M., Peck, E.M., Chauncey, K., Sassaroli, A., Fantini, S., Jacob, R.J.: From brain signals to adaptive interfaces: using fmirs in hci. In: *Brain-Computer Interfaces*, pp. 221–237. Springer (2010)
8. Hayes, G.R.: Knowing by doing: action research as an approach to hci. In: *Ways of Knowing in HCI*, pp. 49–68. Springer (2014)
9. Heger, D., Putze, F., Amma, C., Wand, M., Plotkin, I., Wielatt, T., Schultz, T.: Biosignalsstudio: a flexible framework for biosignal capturing and processing. In: *Annual Conference on Artificial Intelligence*. pp. 33–39. Springer (2010)
10. Kaniusas, E.: Fundamentals of biosignals. In: *Biomedical Signals and Sensors I*, pp. 1–26. Springer (2012)
11. Lauesen, S.: *User Interface Design: A Software Engineering Perspective*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (2005)
12. Marcos, M.C.: Hci (human computer interaction): concepto y desarrollo. *El profesional de la información* 10(6), 4–16 (2001)
13. Martínez, J.G.: *Ingeniería biomédica*. Universidad de Valencia (2004)
14. Moro, R., Berger, P., Bielikova, M.: Towards adaptive brain-computer interfaces: Improving accuracy of detection of event-related potentials. In: *Semantic and Social Media Adaptation and Personalization (SMAP)*. pp. 34–39. IEEE (2017)
15. MyUI: Basic concepts relevant for creating an adaptative myui application (2017), <http://myuidocumentation.clevercherry.com/overview/basic-concepts>
16. Narayanan, S., Koppaka, L., Edala, N., Loritz, D., Daley, R.: Adaptive interface for personalizing information seeking. *CyberPsychology & Behavior* 7(6), 683–688 (2004)

17. Patrick, E.: Intelligent User Interfaces: Introduction and Survey. Delft University of Technology (2003)
18. PMOinformatica: Tecnicas de levantamiento de requerimientos software (2016), <http://www.pmoinformatica.com/2016/08/tecnicas-levantamiento-requerimientos.html>
19. Porta, M.: Human-computer input and output techniques: an analysis of current research and promising applications. *Artificial Intelligence Review* 28(3), 197–226 (2007)
20. Prendinger, H., Mori, J., Mayer, S., Ishizuka, M.: Character-based interfaces adapting to users' autonomic nervous system activity. *IEICE Transactions on Information and Systems* E0-D, 1–6 (2003)
21. Salazar-Ospina, O.M., Rodríguez-Marín, P.A., Ovalle-Carranza, D.A., Duque-Méndez, N.D.: Interfaces adaptativas personalizadas para brindar recomendaciones en repositorios de objetos de aprendizaje. *Tecnura* 21(53), 107–118 (2017)
22. Sánchez-Morillo, D.: Procesado y transmisión de señales biomédicas para el diagnóstico de trastornos y enfermedades del sueño. Ph.D. thesis, Universidad de Cádiz (2008)
23. Schmidt, A.: Biosignals in human-computer interaction. *interactions* 23(1), 76–79 (2015)
24. Shahmohammadi, F., Hosseini, A., King, C.E., Sarrafzadeh, M.: Smartwatch based activity recognition using active learning. In: *Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. pp. 321–329. IEEE (2017)
25. Sitou, W., Spanfelner, B.: Towards requirements engineering for context adaptive systems. In: *Computer Software and Applications Conference. 31st Annual International*. vol. 2, pp. 593–600. IEEE (2007)
26. Vartak, A.: Biosignal Processing Challenges In Emotion Recognition for Adaptive Learning. Ph.D. thesis, University of Central Florida (2010)
27. Voidsa, S., Patterson, D.J., Patel, S.N.: Sensor data streams. In: *Ways of Knowing in HCI*, pp. 291–321. Springer (2014)
28. White, F.: Data fusion lexicon. In: *data fusion subpanel of the joint directors of laboratories technical panel for C3*. pp. 1–13 (1991)
29. Wu, L., Akgunduz, A.: User interface design for artistic expression based on biosignals: An eeg feature extraction method based on weak periodic signal detection. In: *Innovative Design and Manufacturing (ICIDM), Proceedings of the International Conference on*. pp. 5–10. IEEE (2014)

Phase Unwrapping for 3D Object Reconstruction by means of Population-based Metaheuristics

Daniel Duarte-Carrera, Alfonso Rojas-Domínguez, Luis Carlos Padierna

Tecnológico Nacional de México - Instituto Tecnológico de León, León Gto, Mexico
{daducarme,alfonso.rojas,padiernacarlos}@gmail.com

Abstract. Metaheuristics are employed for the solution of the phase unwrapping problem (for 3D object reconstruction) by the *branch cuts* method, posed as an analogous of the traveling salesman problem, which is an NP-hard decision problem. The metaheuristic algorithms carry out a global search for the optimal configuration of the so-called branch cuts which corresponds to a pairing of discontinuities with opposed sign in the wrapped phase map. Three representative algorithms of different metaheuristic families are compared: discrete Particle Swarm Optimization (from bioinspired algorithms), Genetic Algorithms (from evolutionary algorithms) and a novel Estimation of Distribution Algorithm presented in this work that follows a Multinomial distribution. These metaheuristics are comparatively evaluated according to the quality of the solutions achieved, execution time and computational cost, with the aim of building a robust and automated algorithm competitive against traditional methods.

Keywords: Phase unwrapping, Optimization, Estimation of distribution.

1 Introduction

The phase of a signal is often defined within its principal values only, either in $(-\pi, \pi]$ or $(0, 2\pi]$, and it is called true or wrapped phase [1]. In practical applications such as 3D object reconstruction, it is necessary to obtain the phase as a continuous function through a process known as phase unwrapping, which is a technique used to remove the embedded discontinuities in wrapped phase maps [1][2]. The process must detect the 2π discontinuities in the phase and add or subtract 2π an integer number of times to compensate for each discontinuity in subsequent points [3]–[8].

Phase unwrapping algorithms in 2D are most typically divided into two categories: path following or branch cuts methods, and minimum norm methods [2]. The branch cuts method isolates those regions of a phase map that are affected by discontinuities. This is done by the use of barriers or branch cuts, that connect two discontinuity locations, thus achieving path independence [7]. Since its introduction in Goldstein's work in 1988 [9], the branch cuts method has been improved by the incorporation of artificial intelligence techniques (particularly soft computing) [4],[10]. In this work, the branch cuts problem is posed as a computational optimization problem and a comparative evaluation between three types of metaheuristic algorithms is carried out in order to determine their advantages in the solution of said problem. A discrete version of a very popular bio-inspired algorithm, known as Particle Swarm Optimization (d-PSO), is compared against an evolutionary algorithm (Genetic Algorithm) and against a novel

estimation of distribution algorithm: the Multinomial Estimation of Distribution Algorithm (MEDAL). These algorithms are conceptually very different, but whether or not their differences may represent an intrinsic advantage for any of them, is yet to be determined. The goal of this paper is to answer that question.

The rest of this paper is organized as follows: Section 2 provides the required material to understand the formulation of the branch cuts method as an optimization problem. Section 3 briefly presents the different algorithms that are compared. Our experimental methodology and results are reported in Section 4. Finally, Section 5 presents our conclusions and directions for future work.

2 Phase Unwrapping as Optimization Problem

Ghiglia and Pritt explain that there are relatively few inconsistencies along a closed path within a 2D wrapped phase map [2]. These inconsistencies are identified at points where: $\sum_i^M \Delta\psi(p_i) = \pm 2\pi$, where $\Delta\psi(p_i)$ represents the wrapped phase gradient at point $p_i \in \{P\}$ and M is the total amount of points along the path P [2][11][12]. It follows that there are inconsistencies with positive polarity (2π) and with negative polarity (-2π). In 1988 Goldstein used the term *residue* to describe such inconsistencies and described a method where the charge (sign) of each residue must be balanced out by connecting pairs of residues with opposing polarities; this method is known as the branch cuts method [9]. In practice, the residues are computed as the sum of the gradients along a 2×2 path (counterclockwise) given in Eqns. (1) to (4). Whenever said sum gives a positive result, a positive residue exists at position (r, c) ; if the sum is negative, a negative residue is present; if the sum is zero then there is no residue:

$$\Delta\varphi(1) = \text{sign}\{\psi(r+1, c) - \psi(r, c)\}, \quad (1)$$

$$\Delta\varphi(2) = \text{sign}\{\psi(r+1, c+1) - \psi(r+1, c)\}, \quad (2)$$

$$\Delta\varphi(3) = \text{sign}\{\psi(r, c+1) - \psi(r+1, c+1)\}, \quad (3)$$

$$\Delta\varphi(4) = \text{sign}\{\psi(r, c) - \psi(r, c+1)\}. \quad (4)$$

Once the residues have been identified these are connected in pairs of opposing polarity, forming barriers called branch cuts [9]. Then the phase can be unwrapped along any path without touching these barriers. Many different branch cuts configurations can be formed, affecting the complexity of the phase unwrapping process differently. Thus, the phase unwrapping problem is converted into a problem of finding the pairing of residues that produce the optimal branch cuts configuration.

Two branch cuts configurations are shown in Fig. 1. As can be seen, the pairing of residues in Fig. 1b produced four branch cuts that will make phase unwrapping difficult; the barriers are long and badly arranged (crossing each other). In contrast, the pairing in Fig. 1c also produces four branch cuts but their configuration is much more favorable for the phase unwrapping; the barriers are shorter and better distributed. Notice that one of the residues in both configurations has been joined with the border of the phase map. This is acceptable since there is not always an equal amount of positive

and negative residues [7]. An efficient algorithm for this problem must find the pairing of residues that produces the branch cuts configuration with minimum total length. Then the data can be unwrapped by the flood fill algorithm [11].

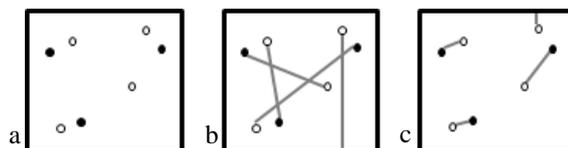


Fig. 1. Two different branch cuts configurations. a.- A set of residues. b.- Unfavorable pairing produces a bad configuration. c.- A different pairing produces an optimal configuration.

The optimization of the branch cuts problem is analogous to a combinatorial problem known as the Traveling Salesman Problem (TSP), which can be summarized as follows [4, 7]: a salesman must visit n cities by means of the shortest possible path; he must visit each city only once and return to the initial city in the end. When the number of cities increases, the TSP cannot be solved in polynomial time since its complexity grows exponentially (it becomes an NP-hard problem). The branch cuts problem is formulated as a TSP problem if the residues are the cities and the sum of the lengths of the branch cuts are the path that the salesman travels [8]. Metaheuristics are effective optimization methods that can be used to tackle this sort of problems.

3 Optimization Metaheuristics for Phase Unwrapping

The different metaheuristic techniques that are compared in this work all share a common codification of the candidate solutions. An individual solution consists in a pairing of residues of opposing polarities (or one residue and one border position). This can always be reorganized as a vector of positive residues and a vector of corresponding negative residues paired with the positive ones. Starting from one solution, new solutions can be generated by keeping the vector of positive residues fixed and changing the position of the negative residues [4]. This is illustrated in Fig. 2 with a small number of residues and wherein some border points, represented by 'B's are included.

All of the metaheuristics discussed herein maintain a so-called population of solutions that they employ to perform their search in the solution space. An initial solution can be found, for example, by application of a simple local search method known as the nearest neighbor method [7]. From this, a population of new solutions is generated automatically by randomly selecting and reordering of k elements in the vector of negative residues. Repeated application of this process creates an initial population.

In order to guide their search for a global optimum, the algorithms evaluate the quality of the individual solutions that are explored. Said quality is quantitatively captured by a so-called fitness function, f . In the case of the branch cuts problem the optimal solution is the branch cuts configuration of minimal total length. Thus, the fitness function is simply the sum of the lengths (distance between each pair of residues in a solution) of branch cuts $((x, y)$ is a residue location), as seen in Eq. (5):

$$f = \sum_i^N [(x_i^+ - x_i^-)^2 + (y_i^+ - y_i^-)^2]^{1/2}. \quad (5)$$

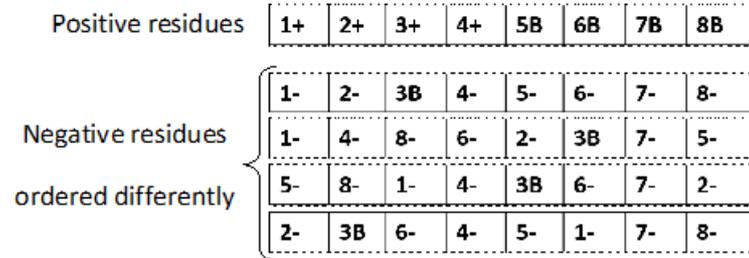


Fig. 2. Codification of solutions. The positive residues are kept in fixed positions; reordering of the negative residues (paired with the positive residues) produces other solutions.

3.1 Bioinspired Algorithm

The particle swarm optimization algorithm (PSO) was developed by Kennedy and Eberhart [12] as a population-based optimization method, inspired on the social behavior of bird flocks. Its objective is to generate increasingly better candidate solutions in an iterative way to reach the optimum [13]. Due to the small number of its parameters, its rapid convergence and its simple implementation, PSO shows better performance than some evolutionary algorithms [13]. PSO is based on the idea of D -dimensional particles moving inside a swarm [4]. The information that the j -th particle uses to move through the search space is the current value of its position U_j , its velocity V_j , its best past position P_j , and the best global position of the swarm P_g .

Table 1. Pseudocode of the PSO algorithm.

PSO algorithm	
1:	Initialize population of D -dimensional particles with random positions and velocities
2:	Start the loop
3:	Evaluate the quality of each particle (solution), according to the fitness function.
4:	Compare each solution with its P_j : if the current value is better, then update P_j and U_j with the current solution values.
5:	Identify the best solution and assign it to P_g .
6:	Update the velocity and the position of each particle: $V_j^{t+1} = (w \times V_j^t) + [C_1 \times Rand \times (P_j^t - U_j^t)] + [C_2 \times Rand \times (P_g^t - U_j^t)]$ $U_j^{t+1} = U_j^t + V_j^{t+1}$
7:	If the stopping criterion is met, stop the cycle.
8:	End the loop

The original process to implement PSO is described in Table 1 [14], where t denotes the iteration number, C_1 and C_2 are non-negative learning factors, the function $Rand$ generates a random number in $(0,1)$ and w is called the inertia factor. A more detailed explanation of these variables is found in [14]. In this work, a discrete variant of the algorithm, termed d-PSO is used because it better fits the branch cuts problem. The modifications are detailed in [4]. In d-PSO the velocity represents a set of permutations

[4]; the permutations modify the position vector, rearranging its values. In this work the adjustment operator is followed in the same way as in [14].

3.2 Evolutionary Algorithm

The genetic algorithm is also a population-based technique inspired by the mechanisms of natural selection, genetics and evolution of living beings that has proved effective, quick and robust in many optimization problems [15]. Members of the population are called chromosomes [16], and they are formed by a set of D genes in D -dimensional space. Chromosomes perform their own adaptation strategies to evolve through the search space and reach the global optimum [7]. A new generation occurs when three operators that update the population are applied: selection, crossover and mutation [16]. The basic steps of the GA are described in Table 2.

Table 2. Pseudocode of the Genetic Algorithm.

Genetic Algorithm	
1:	Initialize chromosome population, probability of reproduction P_r and mutation P_m .
2:	Start the loop
3:	Evaluate the quality of each solution, according to the fitness function.
4:	Use a selection operator to choose two parent chromosomes.
5:	If P_r , Apply the crossover operator. End if
6:	If P_m , Apply the mutation operator. End if
7:	Accept the new solution if its quality is better.
8:	If the stopping criterion is met, stop the cycle.
9:	End the loop

The selection operator chooses a pair of chromosomes for crossover, allowing their genes to pass to the next generation [16]. The crossover operator combines some of the genes of each of the parent chromosomes; both chromosomes are split in the middle and the four parts are combined to form two offspring [7]. Subsequently, the quality of the new chromosomes is measured and only the best is passed on to the next generation [17]. The mutation operator forms a new chromosome through alterations of a chromosome [7]. The heuristic *twors* is used for this purpose [18].

3.3 Multinomial Estimation of Distribution Algorithm

An estimation of distribution algorithm (EDA) is a population-based optimization technique that tracks the statistics of a population of candidate solutions [19]. The search for the global optimum is carried out by creating new and better solutions through these statistics, *i.e.*, recreating the population iteratively and updating the statistics based on the best individuals in each generation [19]. This concept was taken as guidance to design a multinomial distribution EDA for discrete values, since this distribution models the probability of k categories in n trials, which fits well the branch

cuts problem. The steps followed by this Multinomial Estimation of Distribution Algorithm (MEDAL) are shown in Table 3.

Table 3. Pseudocode for the Multinomial Estimation of Distribution Algorithm.

MEDAL algorithm	
1:	Initialize a population of random candidate solutions $\{X_i\}^0, i \in [1, N]$.
2:	Start the loop
3:	Evaluate each solution according to the fitness function.
4:	Select the best M individuals from the population $\{X_i\}^t, (M < N)$.
5:	Estimate a multinomial distribution from the selected M individuals.
6:	Generate a new population $\{X_i\}^{t+1}$ sampling from the multinomial distribution.
7:	If the stopping criterion is met, stop the cycle.
8:	End the loop

During execution of the method, new populations are generated with new probabilities of occurrence per position. The objective is that all individuals converge to one, i.e., the probability of occurrence of a value in a position becomes equal to one (or as high as possible). When this happens, the algorithm ends. Due to the almost null existence of parameters to be tuned, the EDAs are considered agile and efficient algorithms, achieving convergence in a short time despite their susceptibility to get stuck in local optima.

4 Experimental Results and Discussion

The algorithms discussed above were implemented to solve the branch cuts problem and obtain an effective 3D reconstruction of test objects. Each of three test objects is a 512×512 pixels image generated through the MATLAB *peaks* function. For each of these, a wrapped phase map was obtained through the *arctangent* function. The residues in each map were identified by application of Eqs. (1)-(4) as explained in Section 2. Different residue sets were obtained for each map: 1511 residues in the first image (757 positives, 754 negatives); 995 residues in the second image (493 positives, 502 negatives); 1542 residues in the third image (772 positives, 770 negatives). These residues were given to each metaheuristic, together with an initial population of random solutions. This process was repeated 35 times per test image to provide statistical support to our conclusions. Identical initial populations were given to all algorithms in each trial to ensure a fair comparison. The initial populations included 500 solutions each, and the generations per trial were 1000. Tests were performed on an Intel Core-i7 2.40 GHz processor with 8GB of RAM. The test objects (wrapped phase maps), residues, and reconstructions (continuous phase) are shown in Fig. 3. Notice that at this resolution no difference is noticeable between different methods. Therefore, these images are representative of any of the three algorithms tested.

The average total length obtained by each metaheuristic is reported in Table 4, as well as the average elapsed time, function calls (until convergence or completion of the allocated function calls) and mean squared error (MSE) of the reconstructions. The

best results are shown in bold typeface. The stopping criterion was to observe a standard deviation $\sigma = 0.7$ computed over the top 20% of solutions in any generation.

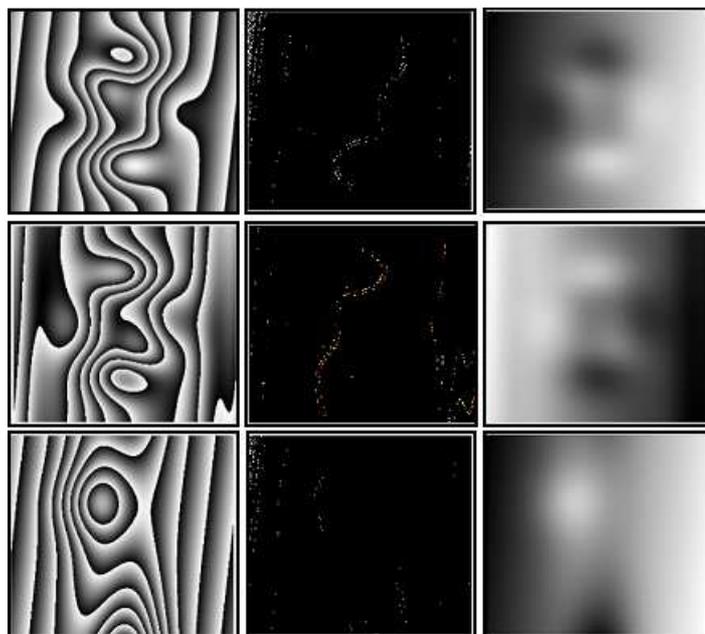


Fig. 3. Wrapped phase maps (left column, 512×512 pixels). Corresponding residues (center column, contrast exaggerated for clarity). Continuous phase obtained by the metaheuristics (right column, at this resolution no difference is noticeable between different methods).

Table 4. Comparison between d-PSO, GA and EDA. Average over 35 trials.

Algorithm	Time (m)	Branch cuts total length	Generations	MSE
First test object				
d-PSO	1.77	2.57 E+03	1000	0.5028
Genetic	2.85	2.58 E+03	423.31	0.8363
EDA	7.11	2.58 E+03	157.11	0.6117
Second test object				
d-PSO	1.32	2.20 E+03	1000	0.3204
Genetic	1.39	2.25 E+03	317.08	0.4573
EDA	2.50	2.28 E+03	132.6	0.5218
Third test object				
d-PSO	1.81	2.02 E+03	1000	0.0665
Genetic	0.68	2.14 E+03	130.08	0.1397
EDA	3.02	2.07 E+03	89.62	0.0434

In Table 4 it can be seen that the three different metaheuristics obtained similar results in terms of average total distance of branch cuts. The lowest average total length was obtained by d-PSO on the three test images. This algorithm also required the lowest execution time in two of the three test objects (first and second). However, the

stopping criterion was never reached by d-PSO, and it consumed the total of allocated generations in every case. In contrast, the GA converged at less than half of the total generations available, and the EDA, at less than a quarter. In other words, the EDA converges much more quickly, but each of its iterations requires more execution time. The behavior of the PSO algorithm is the opposite, and the GA is in the middle of these two extremes.

In order to confirm the statistical significance of the differences observed between the results of the algorithms, and provided that our experiments fit the conditions of a randomized complete block design, the Friedman test (a two-way analysis of variance on ranks) [20] was performed on the results from the 35 experimental trials. In this test, the null hypothesis (h_0) is that the difference in the performance of the algorithms is not statistically significant. Table 5 shows the results of the statistic, the corresponding p-values and the test conclusion at a significance level of $\alpha = 0.05$.

Table 5. Friedman test over 35 trials per test object.

Test object	Statistic	P-value	Result ($\alpha = 0.05$)
Fitness (total length of branch cuts)			
1	1.6	0.4496	Accept h_0
2	10.34	0.0057	Reject h_0
3	27.83	9.05E-07	Reject h_0
Execution time			
1	55.6	8.44E-13	Reject h_0
2	46.69	7.28E-11	Reject h_0
3	51.6	6.24E-12	Reject h_0
Generations (function calls)			
1	62.91	2.17E-14	Reject h_0
2	70	6.30E-16	Reject h_0
3	52.63	3.73E-12	Reject h_0

The Friedman test was applied on our 35 observations of fitness (total branch cuts length); execution time; and generations (function calls), for each of the three test images. As can be seen, the Friedman test rejects the null hypothesis in the majority of the cases. In other words, in the great majority of cases, there is statistical evidence to accept the observed differences between the algorithms reported in Table 4.

Combining the results in Table 4 with the statistical tests in Table 5, important conclusions can be formulated. First, regarding the average fitness measures (third column in Table 4), we consider that the differences are relatively small (and for one test object not statistically significant). Thus, it can be concluded that there is no clear advantage of any method over the rest with respect to the fitness produced; the algorithms are equally effective. In contrast, the differences in the average execution times (second column in Table 4) are quite substantial. As said before, the d-PSO algorithm is the fastest, the EDA the slowest, and the GA is somewhat in the middle. Finally, substantial and significant differences are also observed in the number of Generations required by the algorithms. Here d-PSO is the least efficient, the EDA is the most efficient and the GA is again in the middle. The three algorithms produced good quality solutions (small reconstruction MSE), but it is important to highlight that these measure is not

directly optimized by the algorithms, since their fitness function was formulated in terms of the total branch cuts length only.

5 Conclusions

The problem of branch cuts was formulated in terms of the combinatorial TSP, in order to obtain the benefits gained from years of research on the subject and the use of heuristic techniques to solve it. The success of this novel formulation was demonstrated through the application of three different types of metaheuristics for optimization. The compared techniques are representative of different metaheuristic families: bioinspired, genetic and estimation of distributions algorithms. The Multinomial Estimation of Distribution Algorithm (MEDAL) is a novel formulation of an EDA, created to work with discrete values.

These metaheuristics were applied to solve the branch cuts phase unwrapping problem and these were compared based on their efficiency. The algorithms were tested on three simulated images, demonstrating a fast and efficient reduction of the total branch cuts length and offering better unwrapping results than the initial solutions. Fewer pixels were used as barriers, and smooth continuous phase maps with minor deformities were obtained. All three techniques proved to be efficient, finding better and almost equivalent solutions. Therefore, we consider that they are equally effective in solving this problem.

The fastest algorithm was d-PSO, followed by GA and then MEDAL, with clear and significant differences. On the other hand, MEDAL required substantially fewer generations to converge; on average, it required almost 10 times fewer generations than d-PSO, which did not satisfy the stopping criterion in any of the tests. Thus, with respect to the efficiency of the algorithms, we are faced with contradicting evidence.

Nevertheless, considering that the execution time is dependent on computer characteristics, programming skills, etc., but the number of generations is an objective measure, we conclude in favor of the MEDAL as the most efficient. This conclusion is also influenced by the fact that the MEDAL has no control parameters to be adjusted and therefore it is the most practical for a user to employ.

In future work, the methods studied herein will be employed on real data, combining structured light techniques (to model the objects) with phase shifting to demodulate the phase. Also, the possibility of employing different local techniques to generate different initial solutions will be explored. Based on the results reported, the MEDAL metaheuristic will be considered, because it proved the most efficient on this problem. The experiments presented provide us with the knowledge to make this informed decision, which was not possible prior to the realization of this work. However, extensive testing on real data must be performed before we can conclusively recommend a particular method. The MEDAL metaheuristic will be tested in other applications such as discrete hyper-parameter tuning of classifiers.

Acknowledgements. This work was supported by the National Council of Science and Technology of Mexico (CONACYT) through scholarships 416913 (D. Duarte) and 375524 (L. C. Padierna), and Research Grant CATEDRAS-2598 (A. Rojas).

References

1. Ying, L.: Phase unwrapping. *Wiley Encycl. Biomed. Eng.*, 24, 1–11 (2006)
2. Ghiglia, D.C., Pritt, M.D.: *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. John Wiley & Sons, New Jersey, USA (1998)
3. De Souza, J.C., Oliveira, M.E., dos Santos, P.A.M.: Branch-cut algorithm for optical phase unwrapping. *Opt. Lett.*, 40(15), 3456–3459 (2015)
4. He, W., Cheng, Y., Xia, L., Liu, F.: A new particle swarm optimization-based method for phase unwrapping of MRI data. *Comput. Math. Methods Med.* (2012)
5. Herszterg, I.H.: *2D-Phase Unwrapping via Minimum Spanning Forest with Balance Constraints*. Master's Thesis, Pontificia Universidade Católica do Rio de Janeiro (2015)
6. Huang, Q., Zhou, H., Dong, S., Xu, S.: Parallel Branch-Cut Algorithm Based on Simulated Annealing for Large-Scale Phase Unwrapping. *IEEE Trans. Geosci. Remote Sens.* 53(7), 3833–3846 (2015)
7. Karout, S.A., Gdeisat, M.A., Burton, D.R., Lalor, M.J.: Two-dimensional phase unwrapping using a hybrid genetic algorithm. *Appl. Opt.* 46 (5), 730–743 (2007)
8. Karout, S.A., Gdeisat, M.A., Burton, D.R., Lalor, M.J.: Residue Vector, An Approach to Branch-Cut Placement in Phase Unwrapping: Theoretical Study. *Appl. Opt.* 46(21), 4712–27 (2007)
9. Goldstein, R.M., Zebker, H.A., Werner, C.L.: Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Sci.* 23(4) 713–720 (1988)
10. Wei, Z., Xu, F., Jin, Y.: Phase unwrapping for SAR interferometry based on an ant colony optimization algorithm. *Int. J. Remote Sens.* 293, 711–725(2008)
11. Chen, K., Xi, J., Yu, Y., Chicharo, J.F.: Fast Quality-guided flood-fill phase unwrapping algorithm for three-dimensional Fringe Pattern Profilometry. *Optical Metrology and Inspection for Industrial Applications 7855*, 1–9 (2010)
12. Kennedy J., Eberhart, R.: Particle Swarm Optimization. *Swarm intelligence*, 1(1), 33–57 (2007)
13. Shi, X. H., Liang, Y.C., Lee, H.P., Lu, C., Wang, Q.X.: Particle swarm optimization-based algorithms for TSP and generalized TSP. *Inf. Process. Lett.* 103(5), 169–176 (2007)
14. Poli, R., Kennedy, J., Blackwell, T.: Particle swarm optimization An overview. *Swarm Intell.* 1(1), 33–57 (2007)
15. Holland, J.: *Genetic Algorithms*. *Scientific American* 267(1), 66–73 (1986)
16. Goldberg, D.E.: *Genetic algorithms in search, optimization and machine learning*. Addison Wesley, Reading, MA, USA (1989)
17. Pavez-Lazo B., Soto-Cartes, J.: A deterministic annular crossover genetic algorithm optimisation for the unit commitment problem. *Expert Syst. Appl.* 38(6), 6523–6529 (2011)
18. Abdoun, O., Abouchabaka, J., Tajani, C.: Analyzing the Performance of Mutation Operators to Solve the Travelling Salesman Problem. *International Journal of Emerging Sciences* 2(1), 61–77 (2012)
19. Simons D.: *Evolutionary optimization algorithms*. John Wiley & Sons, New Jersey, USA (2013)
20. Derrac J., García, S., Molina, D., Herrera, F.: A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm Evol. Comput.* 1(1), 3–18 (2011)

Time-Frequency Analysis of EEG Spectrograms using 2-D Gabor Filters for Epileptic Seizure Classification

Ricardo Ramos-Aguilar¹, J. Arturo Olvera-López¹, Ivan Olmos-Pineda¹,
Manuel Martín-Ortíz²

¹ Autonomous University of Puebla, Faculty of Computer Science, Puebla, Mexico
ricramosa1@gmail.com, {aolvera,iolmos}@cs.buap.mx

² Information Tecnology and Communications Department (DCyTIC), National
Laboratory of Supercomputing (LNS - BUAP - CONACYT), Puebla, Mexico
manuel.martin@correo.buap.mx

Abstract. EEG analysis of epileptic seizure is a nonstationary and changing process; these EEGs contain multiple frequencies and use only conventional methods based on frequency or time which limit their analysis. In this paper, texture representation of Gabor filter response based on a spectrogram applying a STFT is proposed to classify epileptic and healthy states. As initial part of our research, energy and entropy were extracted from the Gabor filters response; these features and statistical values were employed to train Support Vector Machines and multilayer Perceptron classifiers.

Keywords: EEG signals, Short time Fourier transform, Spectrogram, Gabor filter, Seizure classification.

1 Introduction

Epilepsy is a neurological disease or disorder distinguished by recurrent unprovoked seizures and caused by the transient and unexpected disturbance, occurring throughout the brain due to hypersynchronous neuronal discharges, which decrease and increase amplitudes of brain activity [7, 21]. Epilepsy affects about 1-2% of the world population and it is divided into two types: focal epilepsy, which involves a part of the cerebral hemisphere and provokes symptoms in some parts of body or related with mental functions; the other kind of epilepsy is the generalized one, which involve the entire brain and produces bilateral motor symptoms, normally resulting in unconsciousness. For this latter kind of epilepsy, there is no specific age at which seizures occur[3].

Electroencephalography (EEG) is a brain monitoring method based on measurements of potentials generated by electrical activity. This test provides evidence of how the brain works over time, which is related to bodily functions such as the pumping of the heart, gland secretion, breathing, internal temperature, etc. EEGs are commonly used by scientists and physicians for analyzing

brain functions and diagnosing neurological disorders, such as brain tumors, head injuries, sleep disorders, dementia, epilepsy, Alzheimer's, seizure disorder, attention deficit disorder, anxiety disorder, fetal alcohol syndrome, and autism, as well as monitoring the effects of anesthesia during surgery [22, 12, 15, 9, 8].

Nowadays, EEG is more used to diagnose and treat neuro-degenerative diseases and abnormalities. Usually, an EEG helps physicians to determine an accurate diagnosis; in neurology, a common application of the EEG is epilepsy detection, since it is a useful and inexpensive tool for showing the underlying manifestations of epilepsy[23]. In people who suffer from epilepsy, their EEG signal shows two categories of abnormal activity: ictal (during an epileptic seizure) and interictal (between seizures)[23, 15]. EEG Analysis for diagnosing epilepsy started in 1970, and since then it has been a changing problem because of EEG's non-stationary features; at present, most problems in seizure detection are related to find events (ictal and interictal) during epileptic seizures [15].

The detection of epileptic seizures in EEG signals has been analyzed using several methods, such as frequency and time-based methods, wavelet transforms, and Gabor filters [21, 15, 14]. However, the nature of the EEG signals is non-stationary and by implementing frequency or time techniques for their analysis[15] presents problems, as the features obtained from these methods do not provide enough information from EEG signals[2]. Time-frequency analysis, however, is a powerful tool that decomposes the signal into both time and frequency, making it possible to analyze non-stationary signals such as EEGs. Using these methods for EEG signals, they can be treated as images and divided in order to extract features from subimages. These techniques have shown good results in terms of accuracy for different applications[20]. Specifically, Short Time Fourier Transform (STFT) is an alternative that has previously used in epileptic seizure analysis by [25] and other recent works with competitive results in classification, where statistical, energy, and other features are extracted from spectrograms [7, 21, 15, 14, 2].

In this initial research work, feature extraction based on a STFT of EEG signals including epileptic seizure activity is presented. This approach extracts features from 2-D Gabor filters images obtained from a spectrogram EEG signal. The feature vector is used as input for training an Artificial Neural Networks (ANN) and Support Vector Machine (SVM). According to our experiments, the proposed method produces acceptable accuracy results. The rest of the paper is organized as follows: Section 2, introduces the methodology and the related theories. Section 3 provides a description of EEG dataset and the experimental analysis. Section 4 presents our conclusions and future work.

2 Methodology

This work presents a methodology for the classification of EEG signals as healthy and as undergoing epileptic seizures, based on first order descriptors obtained from 2D Gabor filters applied to EEG signal spectrograms. Fig. 1 shows the methodology: STFT is first applied to EEG signals to obtain time-frequency

mappings as images, which are then converted into 8-bit gray scale images and divided into five sub-bands corresponding to physiological rhythms on the EEG. 2D Gabor filters are used to extract texture descriptors, which are employed to train SVM and ANN classifiers and to classify EEG signals, showing the results in terms of accuracy.

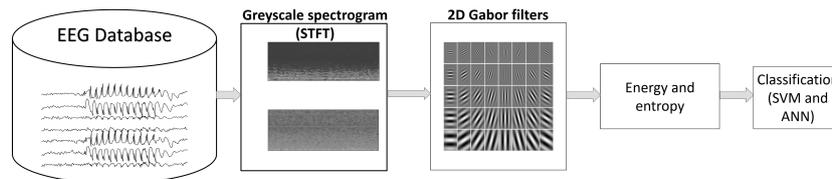


Fig. 1. Proposed methodology: From EEG Databases, spectrograms are computed and Gabor filters are applied to obtain features from the filters response.

2.1 Short Time Fourier Transform

A technique commonly used to analyze electrocardiogram (ECG) and EEG signals is STFT, which maps the signal into a two-dimensional function of frequency and time [18]. To retrieve information from the EEG, STFT is applied to the signal. STFT is employed to partition the EEG signal into several segments of short-time signals by shifting the time window with some overlapping [27], a process called windowing. Depending on the time windowing function $w[n]$, a spectrogram is classified as a narrowband or a wideband. When the time window is short, the STFT will be a wideband and a longer time returns a narrowband spectrogram[5]. The STFT general equation for a signal S is given by Equation (1):

$$S(m, k) = \sum_{n=0}^{N-1} s(m-n)w(n)e^{-j\frac{2\pi}{N}nk}. \quad (1)$$

Where

$k = [0 : K]$ is the K^{th} Fourier coefficient.

$K = N/2$ is the frequency index corresponding to the Nyquist frequency.

$S(m, k)$ indicates the m -index time-frequency(frame) spectrogram.

N =window segment length.

N' =the shifting step of the time window.

$w(n)$ =windowing method of an n point sequence.

N' should be smaller than N in order to produce an overlap between the time windows.

S depends on the window function; in practice, different window shapes are used, such as: Symmetric, Unimodal and Gaussian.

A spectrogram is a plot or visual mapping of the spectrum of frequencies of a signal along the time. This plot has a relation between time resolution and frequency resolution: a large window provides less localization in time and more discrimination in frequency. The window obtains a time-slice of the signal, during which the spectral characteristics are nearly constant [27], and the obtained segments shift the time window with some overlapping. The spectrogram is defined as the magnitude of $S(m, k)$, which is represented as $A(m, k)$ and computed as in equation (2):

$$A(m, k) = \frac{1}{N} |S(m, k)|^2. \quad (2)$$

Spectrogram resolution in time and frequency can change, depending on windowing; a wide window gives better frequency resolution but poor time resolution, while a narrow window produces better time resolution but poor frequency resolution. A visualization with different quality in the spectrogram depends on selecting an appropriate window length and overlapping. Fig. 2 shows the spectrogram of a signal, which is a time-varying spectral representation. The spectrogram layout is usually as follows: the x-axis represents time, the y-axis represents frequency, and the third dimension is amplitude (spectral content) of a frequency-time pair, which is color coded. It is a pseudo-3D plot where the intensity is represented by the z-axis.

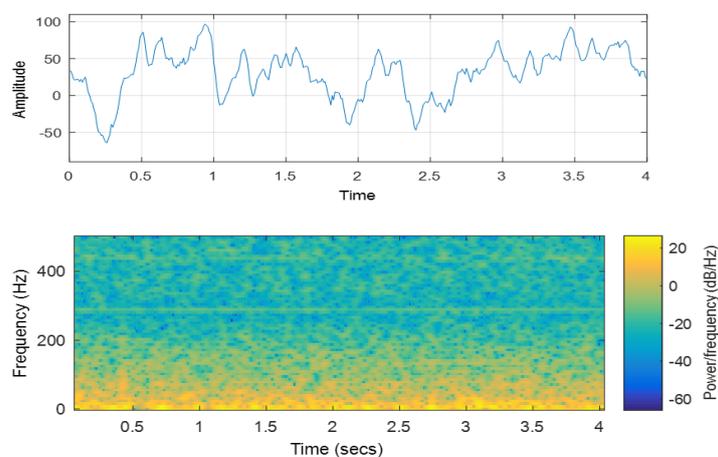


Fig. 2. EEG signal with different amplitude along time and its spectrogram.

2.2 2-D Gabor Filters

Gabor filtering is an option introduced by Daugman for analyzing textures in an image. Gabor filters model the response of simple cortical cells in the visual system and to extract image contours and directional textures [26]; these filters

show an important advantage in their invariance to rotation, scale, translation, photometric disturbances, illumination, and image noise[10]. A 2D Gabor filter is a Gaussian kernel function modulated by a complex sinusoidal plane wave[26], and can be defined as:

$$G(x, y) = \frac{f^2}{\pi\gamma} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp(j2\pi f x' + \phi), \quad (3)$$

Where:

$$x' = x \cos\theta + y \sin\theta, \quad (4)$$

$$y' = -x \sin\theta + y \cos\theta. \quad (5)$$

f is the frequency of the sinusoid.
 θ represents the orientation of the normal stripes in relation to the parallel stripes of the Gabor function.
 ϕ is the phase offset.
 σ is the standard deviation of the Gaussian envelope.
 γ is the spatial aspect ratio that specifies the ellipticity of the support of Gabor function.

2.3 Feature Representation

The Gabor filter response from a spectrogram image is the convolution of Gabor filter with an EEG spectrogram $A(x, y)$. Hence, the response of Gabor filter is given as:

$$R(x, y) = A(x, y) \otimes G(x, y). \quad (6)$$

Where \otimes is the convolution operator.

Using a bank of filters with different parameters, it is possible to obtain features from spectrogram images with different scales and orientations and use them to retrieve local and discriminatory features[6]. A common texture feature is entropy, which is a measure of the amount of information (uncertainty) in a source; in the image field, it measures the randomness of the distribution of the coefficient values over the intensity levels[17]. In this case, the source is an image (spectrogram), which can be seen as a 2D information array[26]. The Shannon entropy is given by:

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i). \quad (7)$$

In the previous equation, $P_r[X = x_i] = p(x_i)$ is the probability mass distribution of the source and n is the number of probabilities $p(x_i)$.

Energy is another frequently used textural feature[16], which measures the uniformity of the intensity level distribution[17, 1] and proposes the energy in an

analysis using Gabor filter texture. The energy is given by the following equation where M and N are the size of the sub-image $a(m,n)$:

$$e(x) = \frac{1}{MN} \sum_{i=1}^M \sum_{i=1}^N |a(m,n)|. \quad (8)$$

2.4 Classification

The energy and entropy features obtained are employed to train SVM and ANN classifiers, and distinguish between epileptic and healthy states. Support Vector Machine is implemented and its results are compared with Artificial Neural Network.

2.5 Support Vector Machines

An analysis technique in supervised learning is Support Vector Machine introduced by Vladimir Vapnik. The generated model outline the training data into different categories separated by a decision boundary. Trying to estimate a function $f : \mathbb{R}^N \rightarrow \{\pm 1\}$ using training data, that is, N -dimensional patterns x_i and class labels y_i , as follows:

$$(x_1, y_1), \dots, (x_l, y_l) \in \mathbb{R}^N \times \{\pm 1\}. \quad (9)$$

such that f will correctly classify new examples (x,y) —that is, $f(x) = y$ for examples (x,y) , which were generated from the same underlying probability distribution $P(x,y)$ as the training data [11]. The decision boundary described by (w,b) , where w is a vector containing the hyperplane parameters and b is an offset. The data is rescaled so that anything on or above the boundary $w^T x + b = 1$ is of one class, and otherwise on or below the boundary is the other class. The in-between distance is called the margin defined as $\frac{2}{\|w\|^2}$. The maximal margin is equivalent to minimize $\|w\|$ which can be achieved through quadratic programming and the optimal hyperplane can be described by: $w = \sum^{\alpha_i} y_i x_i$. Nevertheless, separation task is not always linear due to data representation; hence, SVM transform the data from the input space into a higher dimensional space where is possible separate the data; projecting the data through kernel induced feature space, the kernel function deals with the dot product of the data. The optimal hyperplane can be expressed as equation (10) shows [24]:

$$f(x) = \sum^{\alpha_i} y_i K(x_i, x) + b. \quad (10)$$

2.6 Artificial Neural Networks

McCullock and Pitts proposed a binary threshold unit as a computational model for an artificial method, which computes a weighted sum of n inputs signals, $x_j, =$

1, 2, ..., n, and the output is 1 if the sum is above a specific threshold u . Otherwise, the output is 0. It is defined as:

$$y = \Theta \left(\sum_{j=1}^n w_j x_j - u \right). \quad (11)$$

where $\theta(\cdot)$ is a unit step function at 0, and w_j is the synapse weight associated with the j^{th} input [13]. A classification problem can be defined by the function to use; some functions include step function, linear function and nonlinear sigmoid function. An Artificial Neural Network is constructed combining neurons, which are capable of solving complex real-life problems [24].

In supervised learning, the network is given a desired output for each pattern. Throughout learning, the actual output y generated by the network may not equal the desired output d . A principle of error-correction learning rules is to use the error signal $(d - y)$ to update the connection weights and reduce the error; the perceptron learning rule is based on this error-correction principle. A common network is multilayer perceptron in which each computational unit employs either the thresholding function. This can be form arbitrarily complex decision boundaries and represent any function. Developing a back-propagation learning for determining weights in a multilayer perceptron has made these networks the most populars[13].

3 Experimental Work

3.1 Epilepsy EEG Dataset

The proposed methodology was tested using an open access Epilepsy EEG dataset from the Bonn University [4] with five subsets (Z, O, N, F and S), with each subset containing 100 EEG signals with 4097 samples, recorded at a sampling rate of 173.61 Hz using a 128-channel amplifier system with an average common reference. Sets Z and O were collected from five healthy volunteers. Sets N, F, and S were recorded from five epileptic patients for each set. Records from Set S were collected during seizure activity, while Sets N and F were gathered during seizure-free intervals. Fig. 3 summarizes the Bonn dataset and Fig. 4 shows EEG records from S and Z Sets, which were considered for our experimentation.

3.2 Feature Extraction

The proposed methodology is focused on extracting features from EEG spectrograms through 2D Gabor filters. As noted above, two sets are considered set S (people suffering an epileptic attack) and Z (healthy people); extracted features through Gabor filters are employed to train SVM and ANN classifiers and by means of the extracted patterns the features quality is quantified. To develop

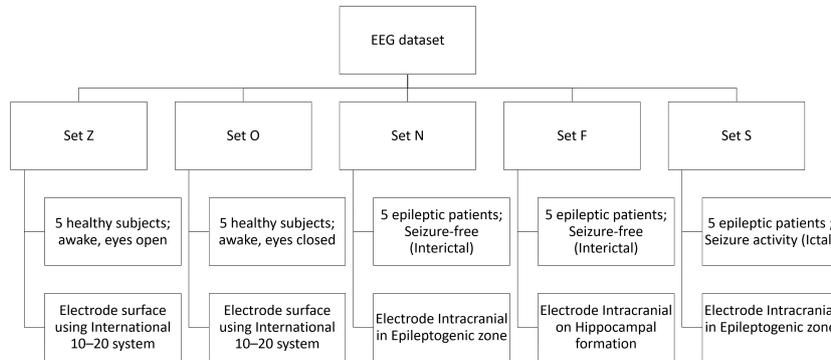


Fig. 3. A summary of Epilepsy EEG dataset.

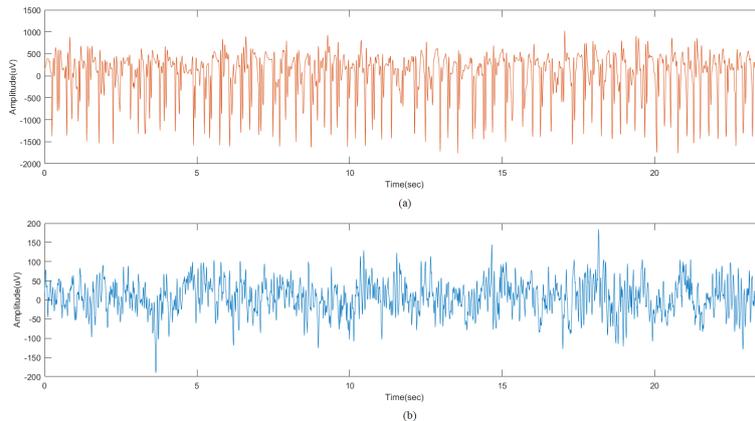


Fig. 4. An example of EEG signals. (a) a person suffering an epileptic attack and (b) a healthy person.

this proposal, Python 3.6 and the packages scikit-learn and scikit-image were used in our experimentation.

As Fig. 5 shows, STFT is first applied to EEG signals from both sets to obtain the spectrograms; the parameters used were the following: a Hanning window with 128 samples was used because of it showed the best result versus other windows, the length of the Fast Fourier Transform $nfft=2000$, and the number of overlapped samples $noverlap=85$. The parameters were selected using

[21] as a reference, because of the results obtained, however, that study did not implement a Hanning window.

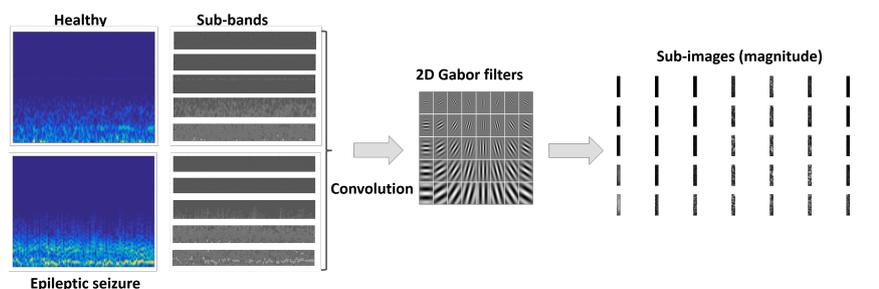


Fig. 5. Generation of the Gabor response from the spectrograms.

The obtained spectrograms were converted into 8-bit gray scale images and then divided into five sub-images related to frequency bands of the rhythms to localize significant structures. This was based on the bands proposed by [21]; other works, such as [7, 19], and [14], proposed dividing the spectrogram; however, they only use four bands, as the gamma band was omitted, since lost information was already taken into account. The ranges employed were: Delta (0-4Hz), Theta (4-8Hz), Alpha (8-12Hz), Beta (12-30Hz) and Gamma (30-50Hz).

Gabor filters were applied to sub-images using five different frequencies and eight orientations; thus, the feature vector was $5 \times 8 \times (tf) \times 5 \times (sb)$; where tf is textural features and sb is sub-bands equal to 5, as shown in Fig 5. To extract the features, magnitudes responses of Gabor filters were considered.

3.3 Experimental Results

Two kinds of textural features (energy and entropy) were obtained and different vectors were generated by gathering the extracted features. To evaluate these features, SVM and with different kernels (linear, polynomial and RBF) and ANN, as different authors reported, such as [7, 15, 21], use this EEG dataset based on spectrograms. The classifiers performance was evaluated using accuracy as a metric, shown as percentage, which is given by the Equation (12); the presented values were obtained from an average of 10 experiments, as 10-fold cross validation was used, selecting the instances in a stratified way:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}. \quad (12)$$

Where:

TP —true positive is a seizure event detected by both the algorithm and the expert neurologist.

TN —true negative is a event identified as non-seizure by the algorithm and the

expert neurologist.

FP—false positive is a seizure event detected by the algorithm, but not by the expert.

FN—false negative is a event identified as seizure by the expert, but not by the algorithm.

In the first experiment, energy was used as feature; hence, the feature vector has a dimension of 200 units. SVM and ANN were trained, obtaining different results, as shown in Table 1. The best accuracy was obtained through SVM-Polynomial, with 98%. SVM-Linear and Perceptron had the same result, although the SVM standard deviation was lower, with 0.024, meaning that 10 experiments of 10 folds were similar; using a perceptron classifier, the number of epochs was changed, using 100 and 500 and the result was the same.

Table 1. Energy feature classification results using SVM and Multilayer perceptron classifiers.

Classifier	Parameters	Accuracy	Standard Deviation
SVM	Polynomial	98%	0.024
SVM	Linear	97%	0.033
SVM	RBF	96%	0.044
Perceptron 1	One hidden layer with 10 units, backpropagation, and 200 epochs.	95%	0.042
Perceptron 2	Three hidden layers with 10, 20 and 10 units, backpropagation, and 200 epochs.	96.5	0.032
Perceptron 3	Two hidden layers with 10 and 10 units, backpropagation, and 200 epochs.	97%	0.024

In addition, entropy was employed as a unique feature and the same classifiers were used as in the first experiment; SVM Polynomial shows the best accuracy, with 97.5%, and the worst was SVM RBF, with 93%; implementing a perceptron was obtained 94% as best accuracy. Nevertheless, energy has the best accuracy between both features. The results are shown in Table 2. Finally, entropy and energy were combined, obtaining a feature vector of 400 units, which were used to train SVM and ANN classifiers. The accuracy was better with a SVM Polynomial, resulting in an accuracy of 98.5% and the worst accuracy was 95.5%, with SVM-RBF. Table 3 shows the results; although these are the best results with polynomial and linear SVM, employing only energy as a feature, the 10 experiments showed less change, with a standard deviation of 0.024.

Some experiments suggest implementing mean and standard deviation of energy as textural features using Gabor filters [16, 1]. However, using standard deviation and mean energy, the accuracy was less than 80% with SVM, while Perceptron with two hidden layers of 20 units reached 95%. Employing stan-

Table 2. Entropy feature classification results using SVM and Multilayer perceptron classifiers.

Classifier	Parameters	Accuracy	Standard Deviation
SVM	Polynomial	97.5%	0.04
SVM	Linear	97%	0.051
SVM	RBF	93%	0.051
Perceptron 4	One hidden layer with 10 units, backpropagation, and 200 epochs.	92%	0.090
Perceptron 5	Two hidden layers with 10 and 10 units, backpropagation, and 200 epochs.	93.5%	0.084
Perceptron 6	Three hidden layers with 10, 50 and 50 units, backpropagation, and 200 epochs.	94%	0.070

Table 3. Classification results with entropy and energy features using SVM and Multilayer perceptron classifiers.

Classifier	Parameters	Accuracy	Standard Deviation
SVM	Polynomial	98.5%	0.032
SVM	Linear	98%	0.04
SVM	RBF	95.5%	0.0437
Perceptron 7	One hidden layer with 10 units, backpropagation, and 200 epochs.	95.5%	0.052
Perceptron 8	Three hidden layers with 10, 40 and 50 units, backpropagation, and 200 epochs.	96.5%	0.032
Perceptron 9	Two hidden layers with 50 and 50 units, backpropagation, and 200 epochs.	97%	0.04

standard deviation and mean to entropy, however, SVM and Perceptron achieved less than 80%. Finally, concatenating four features, the accuracy was 95% for SVM Polynomial, 92.5% for linear SVM, 82.5% for SVM RBF, and 94.5% for Perceptron. Some works as [18, 19, 2, 21] extracted different textural features, in this work one or two kind of features were employed and the results were similar to other approaches.

The obtained results are compared against reported results of other proposals which implement STFT and different features and epilepsy EEG dataset to the same sets S and Z, see table 4. In [7] and [14] five features were obtained with accuracies between 95.78% and 100%, [15] extracted only energy as feature with results between 92.36% and 99.5%. This proposal achieved 98.5% with two features energy and entropy, and this results are better than some experiments.

Table 4. Comparison against other methods.

Reference	Method	Classification
Duque, et al., 2014	Stochastic analysis and cepstrals coefficients.	95.78%-100%,SVM
Kovács, et al., 2014.	MalmquistTakenaka coefficients and statistical features	96.7%, 98.36%,and 99.7% alternating decisiontree
Kumar and Sharma,2015	Energy and PCA.	ANN(99.5%, 99.33% and 92.36%)
Proposed method	Gabor filters, energy and entropy.	98.5% SVM

4 Conclusions

In this paper, an alternative methodology for classifying epileptic seizures is proposed. STFT is used to represent EEG signals as spectrograms and analyze a signal as an image. 2D Gabor filters are employed to extract textural features and SVM and ANN are implemented to classify. Energy and entropy are used as features, however, probably, energy is the predominant feature due to in the most the classifiers, it was better than entropy and the best performance was obtained using SVM Polynomial for any feature. The mean and the standard deviation were not as relevant as in other works; however, with EEG spectrograms, these have been reported showing the behavior of the experiments.

Using the proposed methodology and changing different parameters, such as the length, type and overlapping of the STFT window, accuracy could improve.

From analyzed works, it could be noted that the spectrograms spectral peaks have been analyzed, hence, as future work, it is proposed to analyze the peaks employing clustering techniques to obtain features from the formed clusters and use a method to evaluate features and eliminate some of them non relevant.

Acknowledgments. The first author thanks the support from CONACyT-Mexico through the scholarship number 457637.

References

1. Al-Kadi, O.S.: A gabor filter texture analysis approach for histopathological brain tumor subtype discrimination. *Journal of Science and Technology* 12(22), 1–14 (2017)
2. Alçin, Ö.F., Siuly, S., Bajaj, V., Guo, Y., Şengür, A., Zhang, Y.: Multi-category eeg signal classification developing time-frequency texture features based fisher vector encoding method. *Neurocomputing* 218, 251 – 258 (2016)
3. Ambulkar, N.K., Sharma, S.N.: Detection of epileptic seizure in eeg signals using window width optimized s-transform and artificial neural networks. In: 2015 IEEE Bombay Section Symposium (IBSS). pp. 1–6 (2015)
4. Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P., Elger, C.E.: Indications of nonlinear deterministic and finite-dimensional structures in time

- series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E* 64, 1–8 (2001)
5. Boashash, B.: Heuristic formulation of time-frequency distributions. In: Boashash, B. (ed.) *Time-Frequency Signal Analysis and Processing*, pp. 65–102. Academic Press, Oxford, second edn. (2016)
 6. Dora, L., Agrawal, S., Panda, R., Abraham, A.: An evolutionary single gabor kernel based filter approach to face recognition. *Engineering Applications of Artificial Intelligence* 62, 286 – 301 (2017)
 7. Duque-Muñoz, L., Espinosa-Oviedo, J.J., Castellanos-Dominguez, C.G.: Identification and monitoring of brain activity based on stochastic relevance analysis of short-time eeg rhythms. *BioMedical Engineering OnLine* 13(1), 1–20 (2014)
 8. Gandhi, V.: Chapter 2 - interfacing brain and machine. In: Gandhi, V. (ed.) *Brain-Computer Interfacing for Assistive Robotics*, pp. 7–63 (2015)
 9. Grossi, E., Olivieri, C., Buscema, M.: Diagnosis of autism through eeg processed by advanced computational algorithms: A pilot study. *Computer Methods and Programs in Biomedicine* 142, 73–79 (2017)
 10. Haghighat, M., Zonouz, S., Abdel-Mottaleb, M.: Cloudid: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications* 42(21), 7905 – 7916 (2015)
 11. Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B.: Support vector machines. *IEEE Intelligent Systems and their Applications* 13(4), 18–28 (1998)
 12. Ilyas, M.Z., Saad, P., Ahmad, M.I.: A survey of analysis and classification of eeg signals for brain-computer interfaces. In: 2015 2nd International Conference on Biomedical Engineering (ICoBE). pp. 1–6 (2015)
 13. Jain, A.K., Mao, J., Mohiuddin, K.: Artificial neural networks: A tutorial. *IEEE Computer* 29, 31–44 (1996)
 14. Kovacs, P., Samiee, K., Gabbouj, M.: On application of rational discrete short time fourier transform in epileptic seizure classification. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 5839–5843 (2014)
 15. Kumar, J.S., Bhuvaneshwari, P.: Analysis of electroencephalography (eeg) signals and its categorization study. *Procedia Engineering* 38, 2525–2536 (2012)
 16. Li, W., Mao, K., Zhang, H., Chai, T.: Selection of gabor filters for improved texture feature extraction. In: 2010 IEEE International Conference on Image Processing. pp. 361–364 (2010)
 17. Malik, F., Baharudin, B.: The statistical quantized histogram texture features analysis for image retrieval based on median and laplacian filters in the dct domain. *Int. Arab J. Inf. Technol.* 10, 616–624 (2013)
 18. Mustafa, M., Taib, M.N., Murat, Z.H., Sulaiman, N., Aris, S.A.M.: The analysis of eeg spectrogram image for brainwave balancing application using ann. In: 2011 UkSim 13th International Conference on Computer Modelling and Simulation. pp. 64–68 (2011)
 19. Mustafa, M., Taib, M.N., Murat, Z.H., Hamid, N.H.A.: Gldm texture classification for eeg spectrogram image. In: 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES). pp. 373–376 (2010)
 20. Ridouh, A., Boutana, D., Bourennane, S.: Eeg signals classification based on time frequency analysis. *Journal of Circuits, Systems and Computers* 26(12), 1–26 (2017)
 21. Şengür, A., Guo, Y., Akbulut, Y.: Time-frequency texture descriptors of eeg signals for efficient detection of epileptic seizure. *Brain Informatics* 3(2), 101–108 (2016)

22. Siuly, S., Li, Y., Zhang, Y.: *Electroencephalogram (EEG) and Its Background*, pp. 3–21. Springer International Publishing, Cham (2016)
23. Siuly, S., Li, Y., Zhang, Y.: *Significance of EEG Signals in Medical and Health Research*, pp. 23–41. Springer International Publishing, Cham (2016)
24. Tawfik, N.S., Youssef, S.M., Kholief, M.: A hybrid automated detection of epileptic seizures in eeg records. *Comput. Electr. Eng.* 53(C), 177–190 (2016)
25. Tzallas, A. T., T.M.G., Fotiadis, D.I.: Epileptic seizure detection in eegs using time-frequency analysis. *IEEE Trans. Inf. Technol. Biomed.* 13, 703–710 (2009)
26. Vazquez-Fernandez, E., Dacal-Nieto, A., Martin, F., Torres-Guijarro, S.: Entropy of gabor filtering for image quality assessment. In: Campilho, A., Kamel, M. (eds.) *Image Analysis and Recognition*. pp. 52–61. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
27. Zabidi, A., Mansor, W., Lee, Y., Che Wan Fadzal, C.: Short-time fourier transform analysis of eeg signal generated during imagined writing. In: *2012 International Conference on System Engineering and Technology (ICSET)*. pp. 1–4 (2012)

Emotion Classification of Twitter Data Using an Approach Based on Ranking

Cecilia Reyes-Peña, David Pinto-Avendaño, Darnes Vilariño-Ayala

Benemérita Universidad Autónoma de Puebla, Faculty of Computer Science,
Puebla, Mexico

reyesp.cecilia@gmail.com, {dpinto,darnes}@cs.buap.mx

Abstract. In this work, a model for textual emotion classification based on Ranking technique is presented. The Ranking technique uses the frequencies of words in order to assign a relevance for each in a tweets (Spanish) after calculating the total relevance of the tweet for each classes. The classes are associated to four emotions: happiness, sadness, anger and fear and the highest relevance indicates to which class the tweet belongs. The training and test corpora are created by manually selected key words as references for a crawling tool, both contain manually tagged tweets extracted from Twitter; the training corpus was validated by K-Fold Cross Validation having a 90% of acceptance. The results are compared with Naïve Bayes and Bigrams Probabilities models using precision, recall and F-measure.

Keywords: Emotion classification, Ranking, Twitter, Crawling.

1 Introduction

Every day, it is common the transfer of information through electronic platforms, which are gradually replacing conventional communication services, this means a huge growth in the amount of information available on the web. In this scope, the social networks have become powerful tools for communication in social topics, generating textual and multimedia contents that open up the studies related to the interpretation of emotions that people express. An example of this is Twitter, which is one of the most important social networks and is a huge database related to public opinions on several topics.

Twitter is characterized for being a platform where each user can create or share own short publications (each publication has a limit of 280 characters) or from other user. These publications are known as tweets and could contain one or more hashtag. Hashtags are textual labels, starting with the symbol '#' and when a hashtag is consulted, all tweets which contain the hashtags are shown. If a hastag is very used in a few time, it is consider relevant and is named Trending Topic.

For data extraction in order to create the training and test corpora, Twitter offers an API (Application Programming Interface) and a crawler, which

divide the process into three stages: authentication, crawling and information pre-processing [19].

In this work, we consider four emotion: happiness, sadness, anger and fear, which are opposite to each other (happiness-sadness and anger-fear) [10], and it was necessary to create a corpus that contained tweets associated to each emotion. These corpus is used for training the proposed model and the comparison models: Naïve Bayes and Bigrams Probabilities.

The test results are analyzed with precision, recall and F-measure. Precision is the percentage of the classifier success between all tweets were classified like belong to a class, recall is the percentage of the classifier success between all tweets belong to a class [3], and F-measure is an harmonic between precision and recall and is closer to smaller value of them. The measure most important in this work is Recall because represents the success percentage.

This work is divided as follows: the section 2 contains some related works about emotion classification, the section 3 describes the proposed ranking technique, the Naïve Bayes and Bigrams Probabilities models, in the section 4 the creation of the training and test corpora is introduced, in the section 5 the experimental and measure results are compared among the three models, finally in the section 6 the conclusions and the future works are presented.

2 Related Works

Several projects for natural language processing to identify emotions haven been developed based on different areas as image processing, voice components or textual information. Some related works about textual information retrieval and text classification approach are presented is this section.

There are competition platforms such as SEMEVAL [1] that specializes on semantic similarity systems or tweets classification based on emotion recognition. These systems use manually classified corpus references, which remain static, thus limiting their performance to the new trends of expressions used in Twitter. Therefore, the corpus manual feedback is too late.

Ashequl Qadir et al. [12] propose a tweet tagging system for emotion classification (affect, anger, fear, joy and sadness) using tweets content, patterns and context of the hashtags as base for Bootstrapping technique. Each emotion is associated to five hashtag seeds that represent them. Tweets that cannot be classified in any of the five groups, are tagged by a prefix search process, where the hashtag seeds are used as roots. For feedbacking the system uses the CoTrain technique; in which there are two equal corpora and two models (A and B) with different features, the model A results feedback model B corpus and the model B results feedback model A corpus [11].

Other way for tweets classification is through identification of positive, negative or neutral meaning of each words contained in a tweet and taking into account negative words that can change the sense or context of possible objects, the combination of what is expressed in each word generates a general result [5, 14].

There are approaches for finding figurative emotions that are more complex to analyze such as sarcasm, metaphor and irony by relying on words and hashtags contained in them for identifying the used words (literally, figuratively, etc.)[4]. The approach proposed by Georgios Paltoglou [8] designs a system for emotion classification related to global events on Twitter based on the analysis of either negative or positive polarity changes from the keywords use, leaving behind the popularity indicator through counters.

3 Ranking Technique

In this paper, we propose a model based on information retrieval Ranking technique, that can be defined as a process that assigns a relevance value to each term within a document belonging to a corpus or collection, with the purpose to satisfy a question-answer task[18]. There are two types of classification: simple ranking and aggregation ranking. The simple ranking consists in creating relevance lists that are based on the features of the documents, while the aggregation ranking takes lists of documents already established to form a new one [6].

In this work, in the simple ranking each emotion is used as a document, thus having a collection of four documents. Every document contains tweets and the frequency of each words on each document ($TF_{t,d}$, Term Frequency in the Document). Then, it is necessary to calculate the relevance of each word in the corpus (IDF_t , Inverse Document Frequency of the Term)that can be found with Eq.1, where N is the total documents of corpus, DF_t (Document Frequency of the Term) are the number of documents where the word appears and a base 10 logarithm is used for smoothing the relevance. To get the final score of each word in a document the Eq. 2 is used:

$$IDF_t = \log_{10}(N/DF_t), \quad (1)$$

$$TF - IDF_{t,d} = TF_{t,d} \times IDF_t. \quad (2)$$

When receiving a tweet to classify it, a similar process to a query is used, first each word in the tweet will have a smoothed weight by the \log_{10} of the $TF_{t,q}$ (Term Frequency in the Query) and will be multiplied by the calculated IDF_t before (Eq. 3), the result is a weighted vector which will apply a dot product over $TF - IDF$ values vector related at the same query words for each document. The document with the greatest score is the classification result (see Fig.1):

$$w_{t,q} = \log_{10}(TF_{t,q} \times IDF_t). \quad (3)$$

3.1 Naive Bayes Classifier

Naive Bayes is a popular supervised probabilistic classifier based in Bayes Theorem. It assumes that some feature in particular of a class is independent to the

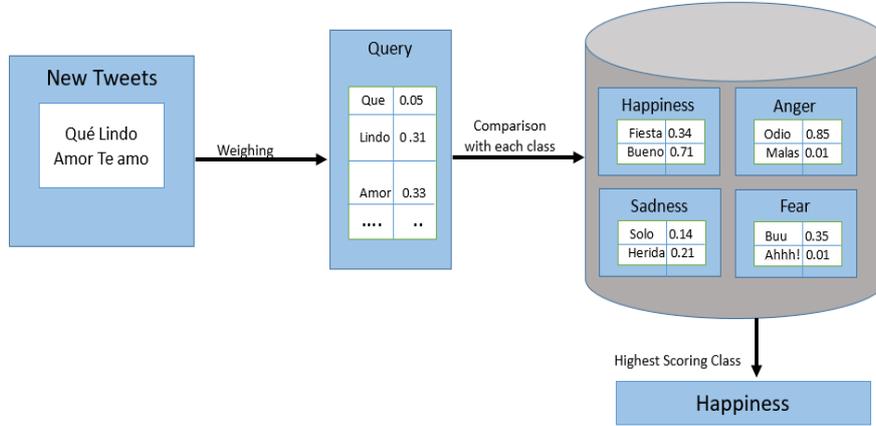


Fig. 1. Ranking Model Diagram.

probability that any other feature belongs or not at the same class [15]. The operation of the classifier lies in calculating the probability of belonging to each feature X in each class C . When the highest value is found, the argument is returned (Eq. 4):

$$\hat{y} = \underset{k \in \{1, \dots, K\}}{\operatorname{argmax}} P(C) \prod_{i=1}^n P(x_i | C_K). \quad (4)$$

3.2 Probability of Bigrams

The N -gram is a series of N consecutive words that belong to a sentence. In the natural language processing, the most used models based on N -grams are commonly unigrams (one word), bigrams (two words) and trigrams (three words) [13]. The degree of N -grams contains more information than grade $N-1$. For the calculation of probabilities in models based on N -grams, the Markov assumption is frequently used, in which it is assumed that the probability of a word depends only on the $N-1$ previous words [2]. Therefore, in a bigram model, the probability of a sentence $P(s)$ made by N words (w_1, w_2, \dots, w_n) is given by the multiplication of the probabilities of each word based the previous one (Eq. 5):

$$P(s) = P(w_1)P(w_2 | w_1) \dots P(w_n | w_{n-1}). \quad (5)$$

Bigrams have a variety of approaches within the scope of natural language processing. One of them is the detection of spelling errors, in which the frequency of each bigram is calculated in a text corpus and those that are infrequent, possibly contain words with spelling errors, it should be noted that in this type of correctors is not possible identify the types of errors that may occur [16]. Another

application of the bigrams is the morphosyntactic labeling from learning having disambiguation benefits based on Markov models [9]. Bigrams have been used in stochastic translation systems by comparing pairs of strings from a source language and a target language [7], and for plagiarism detection [17].

4 Creation of the Corpus

In this work, two sets of tweets manually tagged are used: the training corpus and the test corpus. The training corpus consists of four classes, the tweets were extracted by the Twitter crawler tool, which receives hashtags as keywords associated to each emotion (see Table 1). Each tweet was manually tagged, then the tweets that not express any emotion or expressed contradictory emotions are eliminated. The training corpus size was 105,596 tweets distributed equally for each emotion and for his validation the Naïve Bayes and Bigrams Probabilities classifiers were trained with K -Fold Cross Validation with $K=10$ (the 10% of the corpus for test and 90% for training in ten different iterations) having 90% of acceptance in average.

Table 1. Crawler keywords.

Emotion	Hashtag Associated
Happiness	#Feliz #Bendecido
Sadness	#Tristeza #Triste #RIP #Depression #CorazonRoto
Anger	#ALV #Chingada #HDP #TeOdio
Fear	#Miedo #Terror #TengoMiedo #Pavor

For the test set, some common short words were taken as keywords for the crawler (i.e. a, de, un, la) and the tweets were also manually tagged, we obtain is 234 tweets for the four emotions (127 happiness, 31 sadness, 62 angriest and 14 fear).

5 Experimental Results

About the results, the Ranking model has the highest success number respect to Naïve Bayes and Bigrams Probabilities (see Table 2). According the confusion matrices (see Tables 3-5), the emotion with the highest success percentage is happiness in the ranking model, sadness in Naïve Bayes and fear in Bigrams Probabilities. The emotion with the least success percentage is the anger for Ranking model and Bigrams Probabilities, and happiness for Naïve Bayes.

After applying precision, recall and F-measure (see Table 6), it can be seem that the precision the model with the best average of F-measure is Bigrams Probabilities, Ranking model and the worst performance is from Naïve Bayes.

Table 2. General Results.

Model	Success	Errors
Ranking	129	105
Naïve Bayes	92	142
Bigrams Probabilities	124	110

Table 3. Ranking Model Confusion Matrix.

	Happiness	Sadness	Anger	Fear
Happiness	102	8	10	3
Sadness	8	6	15	1
Anger	7	6	11	0
Fear	10	11	26	10
total	127	31	62	14

In the Ranking model, the small precision values for sadness and fear class show the classifier is tagged many tweets that not belonging for these classes. However, the high value in happiness class indicates that the classifier is hitting in the belonging tweets about the total tagged tweets to these class; in recall, the small values for sadness and anger show there are many tweets that ought to belong to these class and are not assigned in these classes, while the high values in happiness and fear indicate that the classifier is hitting in the tagged tweets between the total the belonging tweets to these class. Given these results, F-measure is low because is closer to the lowest values in all classes.

6 Conclusions

Twitter is a great container of social information, so, the manual classification of tweets that is a difficult task because the people has a different ways of expressing. It is vital to know several approaches and techniques that help in the textual emotion classification as well to know the performance of each to select

Table 4. Naïve Bayes Confusion Matrix.

	Happiness	Sadness	Anger	Fear
Happiness	38	0	1	1
Sadness	53	17	16	1
Anger	9	8	27	2
Fear	27	6	18	10
total	127	31	62	14

Table 5. Bigrams Probabilities Confusion Matrix.

	Happiness	Sadness	Anger	Fear
Happiness	73	5	2	1
Sadness	19	19	15	1
Anger	26	5	21	1
Fear	9	2	24	11
total	127	31	62	14

Table 6. Precision, Recall And F-measure Results.

Model	Measure	Happiness	Sadness	Anger	Fear
Ranking	Precision	0.829	0.2	0.458	0.175
	Recall	0.803	0.19	0.177	0.714
	F-measure	0.816	0.196	0.255	0.281
Naïve Bayes	Precision	0.95	0.195	0.586	0.163
	Recall	0.299	0.548	0.435	0.714
	F-measure	0.455	0.288	0.5	0.266
Bigrams Probabilities	Precision	0.901	0.351	0.396	0.239
	Recall	0.574	0.612	0.338	0.785
	F-measure	0.701	0.447	0.365	0.366

the one that provides a better result. This work proposes a new approach about the information retrieval ranking technique, which taking the highest relevant score as a classification result.

Being recall the most important measure in this work the classifier Bigrams Probabilities has the highest average for it despite of the F-measure's average of all classifiers is under the 50% of success. It shows the bigrams approach offers better results.

In future work, it is necessary to modify the Ranking model to get better results, adding other classifiers in order to apply classifiers ensemble for selecting a tweets set and based on it the implementation of automatic corpus feedback technique (Bootstrapping).

References

1. International workshop on semantic evaluation (2016), <http://alt.qcri.org/semEval2017/index.php?id=tasks>
2. Basharin, G.P., Langville, A.N., Naumov, V.A.: The life and work of aa markov. Linear algebra and its applications 386, 3–26 (2004)
3. Feldman, R., Sanger, J.: The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge university press (2007)

4. Ghosh, A., Li, G., Veale, T., Rosso, P., Shutova, E., Barnden, J., Reyes, A.: Semeval-2015 task 11: Sentiment analysis of figurative language in twitter. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). pp. 470–478 (2015)
5. Gonzalez-Marron, D., Mejia-Guzman, D., Gonzalez-Mendoza, M., Enciso-Gonzalez, A.: Explotacion de informacion de redes sociales de microbloggs utilizando analisis de sentimientos y tecnicas propuestas por el big data. *Tecnologías de la Información* 2(2), 90–103 (2015)
6. Li, H.: Learning to rank for information retrieval and natural language processing. *Synthesis Lectures on Human Language Technologies* 4(1), 1–113 (2011)
7. Mariño, J.B., Crego, J.M., Lambert, P., Banchs, R., de Gispert, A., R Fonollosa, J.A., Costa-Jussá, M.R.: Modelo estocástico de traducción basado en n-gramas de tuplas bilingües y combinación log-lineal de características. *Procesamiento del Lenguaje Natural* (35) (2005)
8. Paltoglou, G.: Sentiment-based event detection in t witter. *Journal of the Association for Information Science and Technology* 67(7), 1576–1587 (2016)
9. Pla, F., Molina, A., Prieto, N.: Evaluación de un etiquetador morfosintáctico basado en bigramas especializados para el castellano. *Procesamiento del lenguaje natural* 27 (2001)
10. Plutchik, R.: The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist* 89(4), 344–350 (2001)
11. Qadir, A., Riloff, E.: Bootstrapped learning of emotion hashtags# hashtags4you. In: Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis. pp. 2–11 (2013)
12. Qadir, A., Riloff, E.: Learning emotion indicators from tweets: Hashtags, hashtag patterns, and phrases. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1203–1209 (2014)
13. Rodríguez, H.: Lingüística y estadística,¿ incompatibles. *Tecnologías del texto y del habla* 72, 89–115 (2004)
14. Rosenthal, S., Nakov, P., Kiritchenko, S., Mohammad, S., Ritter, A., Stoyanov, V.: Semeval-2015 task 10: Sentiment analysis in twitter. In: Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015). pp. 451–463 (2015)
15. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited, (2016)
16. San Mateo, A.: Un corpus de bigramas utilizado como corrector ortográfico y gramatical destinado a hablantes nativos de español. *Revista signos* 49(90), 94–118 (2016)
17. Torrejón, D.A.R., Ramos, J.M.M.: Detección de plagio en documentos. sistema externo monolingüe de altas prestaciones basado en n-gramas contextuales. *Procesamiento del lenguaje natural* 45, 49–57 (2010)
18. Trotman, A.: Learning to rank. *Information Retrieval* 8(3), 359–381 (2005)
19. Zheng, Z., Alonso-Berrocal, J.L., G Figuerola, C.: Análisis cibernético y visual de Twitter. Departamento de Informática y Automática. Facultad de Ciencias. Universidad de Salamanca (2013)

Automatic Segmentation in Breast Thermographic Images Based on Local Pattern Variations

Daniel Sánchez-Ruiz, Iván Olmos-Pineda, J. Arturo Olvera-López

Autonomous University of Puebla, Faculty of Computer Science, Puebla, Mexico
daniel.sanchez.4712@gmail.com, {iolmos,aolvera}@cs.buap.mx

Abstract. In order to detect any possible anomaly in the region of breast, the use of thermal images has experienced a considerable workload of research in the recent years, this due to the promising results of this technique. One of the principal tasks in this process is the segmentation of the Region of Interest (ROI), but this task is difficult, most of the proposed techniques perform a manual or semi-automatic process to extract it. In this paper, we propose an alternative technique to detect the ROI in thermal breast images. In the proposed technique, we focus on the higher temperature areas in the image. We apply local contrast enhancement, group higher temperature regions, spline cubic interpolation and statistical operations as a part of the method. The achieved results are competitive with the state of the art showing a new alternative to accomplish the automatic segmentation of thermal breast images.

Keywords: Automatic segmentation, Contrast enhancement, Image processing, Thermography.

1 Introduction

According to the World Health Organization from the last decade, cancer diseases are the second dealing cause of global death, by 2015 cancer accounted 8.8 millions of deaths. There are several types of cancer but perhaps one of the most dangerous and common is breast cancer, which has occasioned 571 000 deaths [1]. Cancer mortality can be reduced if cases are detected and treated early [1]. Breast cancer usually appears in ducts; tubes that carry milk to the nipple, and lobules; glands that produce milk. Some works have reported that the growth rate of a tumor is proportional to its temperature [17]. There are a lot of research in this field, one of the current investigations to detect anomalies in early stages is the use of digital thermal images which take into account patterns based on temperature changes.

There exist several procedures in order to diagnose breast cancer such as Mammograms, Ultrasounds, Magnetic Resonance Imaging (MRI) and others [2, 3, 4]. The mammogram is a widely used study, but the problem is the limited range of population whom are candidates (young women are not suitable), the invasive and even hurtful procedure, the expensive devices utilized, and finally it requires repeated exposure to radiation (who is nocive for human health [5]). The ultrasound and the MRI

are commonly performed as the previous step of a mammography, but in spite of its good results, none of these studies are 100% accurate, the only study that achieves that precision is the biopsy.

Infrared images do not use ionizing radiation, venous access, or other invasive procedures. The infrared image presents physiological information of normal and abnormal functioning of the vascular system, sensorial and sympathetic nervous system, and inflammatory processes [6, 7]. The combination of mammographies and thermography, allows the achievement of a high degree of specificity and sensibility on such diagnosis [8, 10]. Moreover, thermography is very useful for detecting non palpable breast cancer in earlier stages [6]. Nevertheless it has to be clear and well understood that thermography by itself cannot replace any of the other studies, it is a complement of all of them.

In the digital images analysis process, an important and difficult task to achieve is the ROI segmentation, which allows the delimitation of the data to be analyzed. In the case of thermal breast images, the ROI must include all the breast tissue and near ganglion groups since cancerous cells usually appear in the glands that produce milk and the ducts that carry it to the nipples [17]. An accurate segmentation of ROI from medical thermal images still remains as an open problem. Most authors perform semi-automatic or manual ROI extraction because it is a hard task due to inherent limitations of thermal images such as the absence of clear edges, low contrast nature and low signal to noise ratio [9].

There are various techniques presented in reviews that perform automatic segmentation [9, 11, 18]. For example Marques et al. [12] presented an automatic segmentation method which consists of image processing techniques such as thresholding, clustering, edge detection and refinement among others. Villalobos-Montiel et al. [13] used Canny edge detection with automatic threshold and symmetry inspection, besides Hough transform and Active Contours are used to provide a better approximation to the ROI. Hinojosa et al. [19] proposed the use of thresholding techniques as objective functions in evolutionary algorithms, in particular the HS (Harmony Search) algorithm provided excellent results. Ali et al. [20] used the distance between camera and patient as a relevant parameter, they assumed the ROI is always in the same region so they defined fixed regions to obtain it, but their approach fails many times due to the lack of dynamic in their method. Sathish et al. [21] have focused on shape features of the breast and polynomial curve fitting, also a statistical test is performed to evaluate the results who are competitive with the related work. Hankare et al. [22] proposed a method who performs a color space transformation from RGB to a $L^*a^*b^*$ color space, then a k-means classification with an Euclidean distance is applied, finally the cluster with the biggest temperatures is segmented. Prakash et al. [23] presented three segmentation techniques K-Means, Fuzzy C-Means and Gaussian Mixture Model – Expectation Maximization, in this paper also the color space transformation was important to achieve good results. Sayed et al. [24] proposed a method using bio-inspired swarm techniques forming clusters looking for the most optimal, the FA (Firefly Algorithm) algorithm was the best swarm version that obtained the highest accuracy, sensitivity, precision and specificity. Mejia et al. [25] used morphological operators in order to enhance the ROI area, then thresholding and the Extended-

minima transform are performed to search neighborhoods connected who delimited the ROI region. Garduño-Ramón et al. [26] employed morphological watershed operator to help to locate a possible tumor, thresholding operations and polynomial curve fitting were also performed.

However, most techniques fail to perform a correct segmentation in images where the borders in the ROI are unclear due to the different physical characteristics of each person besides many focus only in the tumor region and not in the whole area of the breast and with this approach a common problem is the irregular shape of malignant tumors.

In this paper we present a technique that performs an automatic ROI segmentation for thermal breast images. This consists of four main stages, the first stage extracts the background from the image, Otsu's method [15] is applied; the second stage finds the bottom boundaries of the ROI, to achieve this, automatic thresholding methods are used, local contrast enhancement, clustering of regions that represent the highest temperatures and polynomial curve fitting through interpolation methods are also performed; the third stage is focused on detect the upper boundaries, statistical analysis in the area of lateral borders is applied, finally in the fourth stage the segmentation of each breast and also some correction operations are performed.

The organization of the paper is the following, in section 2 the methodology for the proposal is presented, in section 3 results and discussion are presented, finally in section 4 the conclusions are drawn.

2 Breast Image Segmentation Methods

In order to perform an automatic segmentation of the ROI in a thermal breast image, a technique that combines different methods over distinct stages is proposed. The technique consists of four stages: background segmentation, detection of bottom boundaries, detection of upper boundaries and segmentation of each breast, Fig. 1 shows the complete process and the result of every stage. In following subsections all the details of each stage are described.

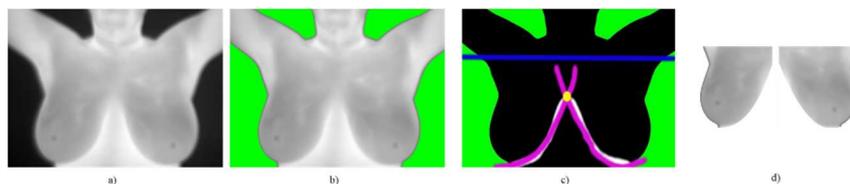


Fig. 1. a) Input image, b) Background Segmentation, c) Detection of upper and bottom boundaries, d) Segmentation of each breast.

2.1 Database Description

The DMR (Database for Mastology Research) Database [14] is a public available breast thermogram database. The database contains breast thermograms of 287 sub-

jects out of which, the thermograms of 47 subjects are labeled as ‘Sick’ and remaining thermograms are labeled as ‘Healthy’. For acquisition of the thermograms, a FLIR SC-620 Thermal Camera with a spatial resolution of 640 x 480 pixels was used. All the images are in RGB format but the information on each channel is the same, they are in grayscale format.

2.2 Background Segmentation

With the aim to segment the body from the background, it is implemented a segmentation based on color intensities. Fig. 2 shows in the histogram, the contrast in a thermal image between the body and the background, clearly it can be noticed two regions separated by a valley. The region on the left represents the darker color values (background) and the region of the right represents the brighter color values (body) in the image. In order to separate both regions we found a threshold value between them. Each region can be represented as a class of color, and a common technique for automatically find a threshold value between two color intensity classes is the Otsu’s method [15], so the first step is apply this method to the input image.

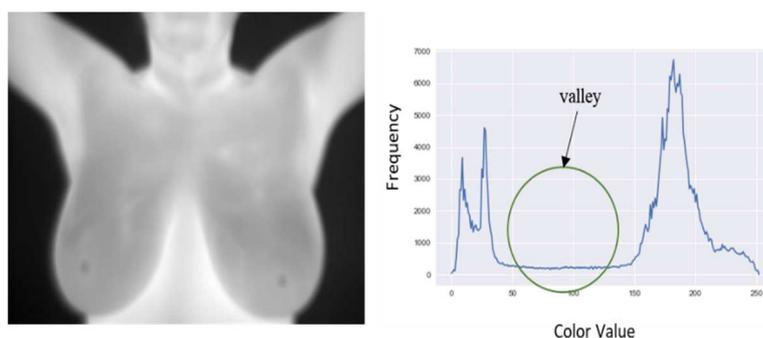


Fig. 2. Grayscale image histogram showing two regions and a valley between them.

With the threshold value found by Otsu's method, all the values which are under the threshold are considered as part of the background. With the objective of detect easily this region, a control color value is considered and applied to the image. Equation 1 shows the applied rule, where x is the input color value, cc is the control color value and th is the threshold value found:

$$f_{th}(x) = \begin{cases} cc, & x \leq th \\ x, & x > th \end{cases} \quad (1)$$

In Fig. 1b the result of applying the background separation is shown. In the segmentation process in some cases, some parts of the body such as nipples are modified due the threshold value, a solution to this affected regions is presented in section 4.

2.3 Detection of Bottom Boundaries of the Breasts

The next stage is one of the most difficult in the whole process, the detection of bottom boundaries. In this work we propose an approach based on local patterns such as variations in color intensities. The first step of this stage is automatically find a threshold value that it can separate the regions related with the highest temperatures in the body from the rest. In thermal images the highest temperatures are represented with brighter values and the lower temperatures with darker values. This is relevant for the database images because the regions with the highest temperatures are present in armpit, neck, and under the breasts; the latter region is what we want to identify. Fig. 3 shows bottom boundaries of the breasts where the temperatures are higher in comparison with others regions in the body, these temperatures are represented with brighter values.

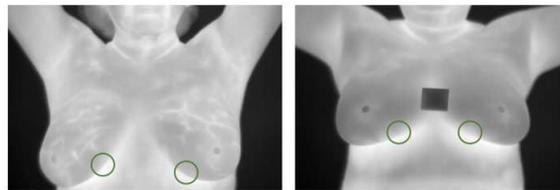


Fig. 3. Higher temperatures in bottom boundaries of the breasts.

In order to face this problem a local analysis is carried out. The image is divided in three sub regions from up to bottom, and in other three sub regions from left to the right. After this process the image has nine sub regions. The principal sub regions to analyze are the middle column in the second row, and the left and right sub region in the last row, because these areas contain the information related with the bottom boundaries of the breasts in most of the images, only in a few images this may not be true. The division of the image is showed in the Fig. 4. After identify this sub regions the technique discard the rest of the sub regions to the following operations.

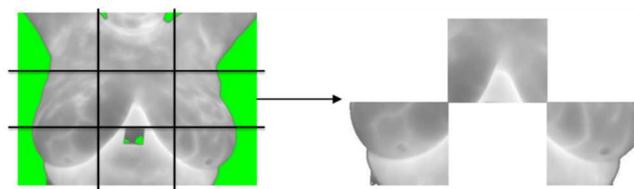


Fig. 4. Image divide into nine regions and the selected images.

In every image of the chosen sub regions an accumulative contrast enhancement [16] is performed, so the brighter color values are separated from the dark color values, this is important in order to discard the lower temperatures values of the breasts because these values are considered as noise in this stage and they can affect the posterior analysis. Fig. 5a shows the result and the histogram of the original image and after the contrast enhancement is applied in one of the selected sub regions.

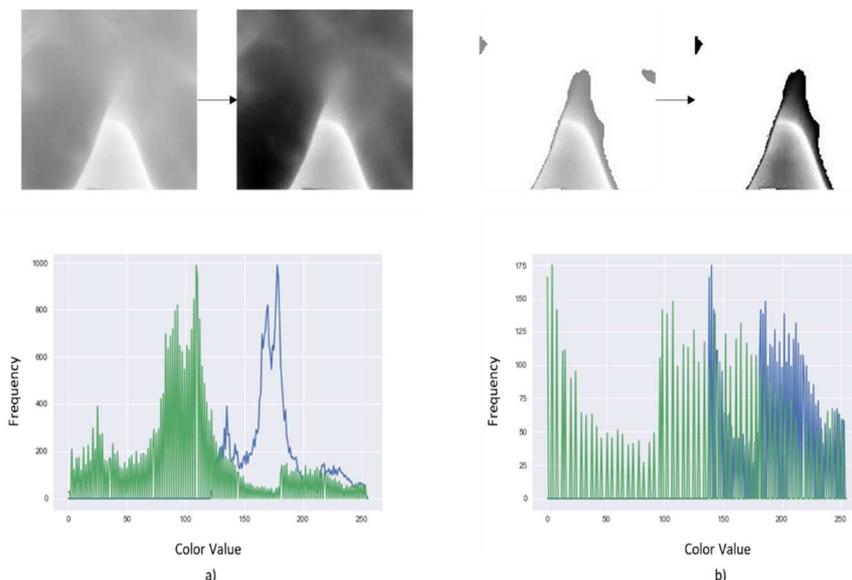


Fig. 5. a) Original image and contrast enhancement image and its histograms (blue for the original and green for the altered image), b) Image segmented and contrast enhancement and its histograms (blue for the original and green for the altered image).

With the aim to remove the darkest color values, Otsu’s method is applied. Any color value that is over the threshold value found is segmented. At this moment it has been removed many dark values as is showed in the left image in Fig. 5b, and also it can be notice that the region under the breast it remains. In the segmented region a second contrast enhancement is applied to enlarge the difference between the bright and dark color values. The Fig. 5b shows the result of second contrast enhancement, it can be noticed how it was enhanced the region of the bottom boundaries of the breasts in the image. Now the boundaries are better than in the input image.

Otsu’s method is performed one last time to obtain a threshold value who it can segment the brightest color values in the local sub region, this returned value is stored in a list. After this procedure is applied for every one of the three sub regions, three local threshold values are found, with these values the mean is calculated, this is considered as the global threshold value, lastly with this value a binarization is applied to the input image of this stage. Fig 6 shows a diagram of the process to follow in the chosen sub regions to automatically find the threshold value who segments the highest temperatures regions from the rest.

Fig. 7a shows in white color the regions of the highest temperatures present in the body after the binarization. As it was expected, these regions are in the neck, armpits, under the breasts and in some small regions around the body. Also the biggest region is under the breasts, so the procedure focus in locate it and discard the rest of them. The regions of the upper half body are discarded because they are not relevant to this stage, so they are set it to a black color.

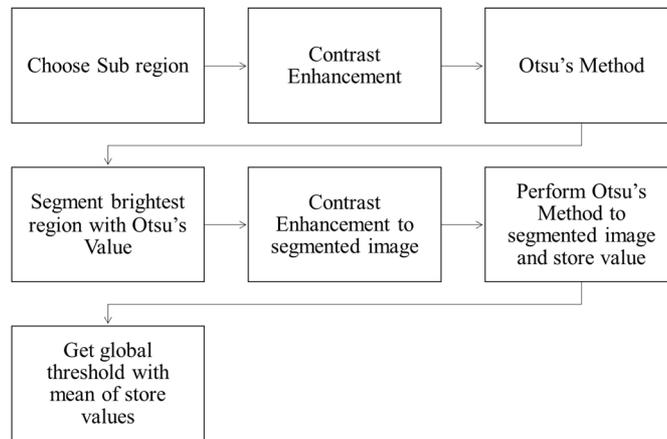


Fig. 6. Diagram of the process to automatically find the threshold value who segments the highest temperatures, the first six steps are performed for every sub region, the last one is performed a single time.

Then we form clusters, regions that are composed by white points that are neighbors between them, for every point is stored its coordinates. A cluster can be view as a list of points. The clusters are stored in another list, and sorted from the region with more elements to the region with fewer elements. The technique only analyses the two biggest regions, in which it is checked a condition between them, if the proportion of the second region with more elements is at least 34% (this value keeps the most relevant data and we found it through empirical experiments over all the images), then this region is considered part of the boundary, but if this condition is not reached, only the region with more elements is considered, all the remaining regions are discarded and set it to a black color. The Fig. 7b illustrates the final result of this operation.

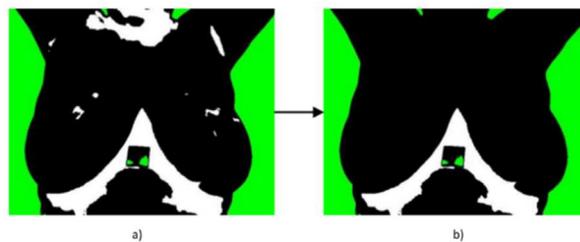


Fig. 7. a) Image binarized with the threshold found, b) Final result after removed the undesired regions.

Finally a simple process of thinning is applied. In a loop over the image from the left most column to the right most column, and from the first row to the last, when a white color value is detected, a control flag is turn on, while this flag is on any other white color value detected in the same column is set it to a black color value, when

the loop pass to another column this flag is turn off until is detected another white color value. The final result are two curves which represent the bottom boundaries of each breast.

Then a spline cubic interpolation by the Lagrange's method is performed for every breast with the aim to fit the detected curves, the method return a polynomial that it can be used to extrapolate the curve to any other point. Finally with the polynomials, the intersection point between the curves can be calculated equalizing them to zero and solving the equation. The final result is shown in Fig. 8.



Fig. 8. Thinned line in white, curve fitting in purple and intersection point in yellow.

2.4 Detection of Upper Boundaries

The next task is to find the upper boundaries. For this step it is used the information related to each boundary side and the information of the previous stage. Along with this data some statistical operations are performed. The coordinates in the x and y axis where exists a transition from the background to the body are saved in a list, one list for each side. The lists are generated from down to up of the image, so the first points correspond to coordinates close to the breasts and the latest are near of the arms in the image. Then is calculated the mean μ and deviation standard σ of the x coordinate values for each list as is showed in equation 2 and 3:

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i, \quad (2)$$

$$\sigma = \sqrt{\sigma^2} \text{ and } \sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2. \quad (3)$$

Where X_i is the x coordinate value and n is the total points in the list. Due to the body posture in the images, the mean is close to the armpit and arms, this happens because the majority of the x coordinate values are present in this area. The standard deviation give a plus/minus range of how much the values are spread from the mean, the technique find the first value outside this range. As it is known the order of the list (down to up of image), in a loop over all of the points in each list, it is going to be found the first point where its x coordinate value is greater than the sum of the mean plus/minus (depending of the body side) deviation standard, this point is stored. This analysis is made for each side. The points found are usually different and only one point is considered, a comparison is performed between them and the one with the highest y coordinate value is selected. Besides that, it needs to be checked if the selected point is inside the area delimited by the point of intersection between the breasts found in the

previous stage and the first row of the image, this is done with the aim to avoid bad results. Fig. 9a shows an example of the stored points of each side, the mean and standard deviation calculated and Fig. 9b shows the final result of this stage.

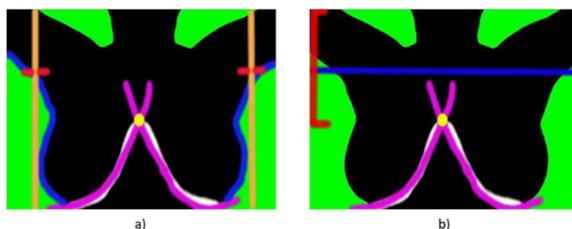


Fig. 9. a) Points over the blue region are stored in the described lists, the mean in the x coordinates is close the orange lines and the standard deviation is a plus/minus range represented with the red lines, b) In red is shown the check range where the upper boundary needs to be, in blue the final upper boundary detected.

2.5 Correction and Segmentation

Finally the segmentation of each breast is performed, to achieve this it is used the data of the boundaries and the point of intersection of the previous stages. All the points inside the upper, bottom and lateral boundaries are considered, and the point of intersection between the breasts is used as middle point, finally two images are created and the background color is removed as a final step.

The input image of this stage is the output image of the first stage, this image is very useful because the lateral boundaries are known due the control color of background and all the original values in the body are present in the rest of the image. Sometimes a few parts of the body have very dark values that are wrongly set it to the background color, so this represents a loss of information. If this case happens a correction must be applied with the image resulting of the first stage and the input image (original image). This is fixed with a procedure that visits all the neighbors in a window of 3x3 for a central point. This procedure checks if the value for the central point is the same of the most repetitive value of its neighbors, if the condition fails, the value for the central point is restored to the original value. Fig. 10 shows a corrupted image, and the result of the correction.

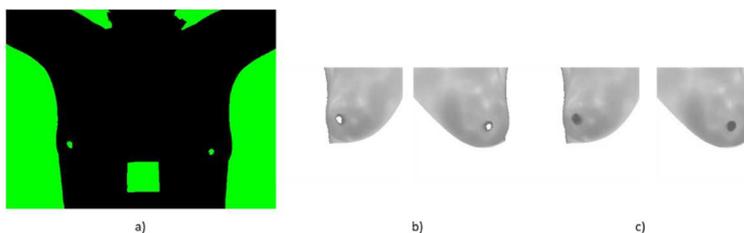


Fig. 10. a) Image corrupted by the procedure of the first stage, b) Segmented image with loss of information, c) Final Image with original values recovered.

3 Experiments and Discussion

The proposed method was tested on 207 images from the DMR database [14] in order to evaluate it. The development of the technique was realized with Python, most of the project was an original development but some packages were used, such as scipy [27], numpy [28] and scikit-image [29].

Some results are showed in Fig. 11, the image of the first row shows an example of a case with clear color transitions under the breast and in the body, the result proves that the method achieved a very good result; the image of the second row presents a case where one of the breast has been part of a medical procedure, in spite of this, the method approximate very accurately the region of both breasts; in the final row it is presented an image where the color transition in the boundaries under the breast are not clear, nevertheless the result is also pretty good. All this results showed that the proposed technique works well in many possible scenarios.

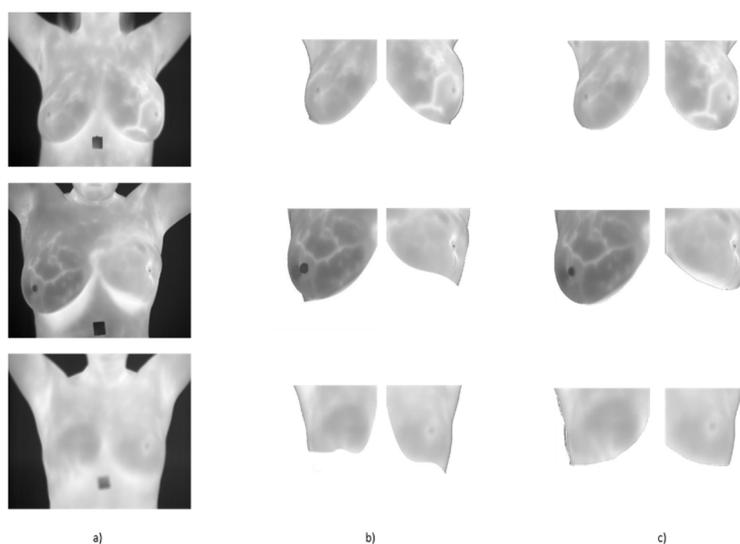


Fig. 11. a) Input image, b) Segmented ROI by the proposed technique, c) Segmented ROI in the Ground Truth.

After performed tests in all the images around 85% were segmented with a result from acceptable to accurate. It was defined as an acceptable result any outcome where the ROI does not suffer any loss of information, even if some extra information is present in the image, e.g. small areas under the breasts that are not part of the ROI. The acceptable images are very useful because they introduce only a small amount of noise (data not belonging to the ROI), so this images can be used in future analysis such as pattern recognition without seriously affect the results. Some examples of different results are showed in Fig. 12. A summary of each kind of image is presented in Table 1.

In order to evaluate the obtained results, the segmented images by the proposed technique and a Ground Truth (elaborated by ourselves) of manual segmented images were used. Each image was compared with its counterpart in the Ground Truth and the difference between each one was calculated as a percentage. Fig. 11b shows some examples of the Ground Truth. Finally, the average of all percentages were calculated as an accuracy measure. Comparing with the research of Marques et al. [12] and Villalobos-Montiel et al. [13] who used the same database, the outcome is competitive, the results are listed in Table 2.

Table 1. Analysis of the segmentation process result.

Result	Total	Percentage
Accurate	160	77.3%
Acceptable	15	7.2%
Poorly	32	15.5%

Table 2. Comparison in segmentation accuracy.

Author	Accuracy
Marques et al.	96.0%
Villalobos-Montiel et al.	98.7%
Proposed technique	94.0%

The main contribution of this work is it only takes into account simple operations such as contrast enhancement, local analysis and statistical operators in order to find the upper and lower boundaries. Obviously this techniques will not work with other databases because the acquisition protocol of the images might be different, and an important fact for the method, at least for the step of detecting the upper boundaries in the ROI is the applied protocol in the DMR database, but a brief analysis and few modifications are necessary to achieve the same results with a new data set.

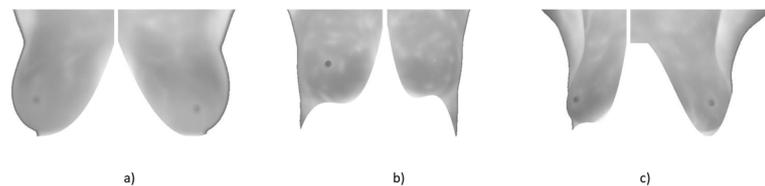


Fig. 12. a) Remarkable segmented image, b) Acceptable segmented image, c) Poorly segmented image.

4 Conclusions

This paper presented a technique for automatic segmentation in breast thermal images. The ROI extraction in medical domains is an important step in the development of systems for medical diagnosis based on digital images, because an accurate segmentation allows a better analysis of the most relevant information in the data. It was shown through the methodology and the performed experiments that this is not an easy and trivial task.

The proposed approach is based on local analysis with the aim to avoid global noise. Intrinsic characteristics in the image are used to identify borders, such as body posture or temperatures related with color values.

The technique showed good results even in images where the borders of the ROI are unclear, in spite of this, there are some errors in images with low contrast in the region of the breasts, amorphous breasts and in images with too many regions of high temperatures around all the body. These cases are an opportunity to keep working in the technique and improve the results. As a future work, the second stage of this project will be addressed, the stage related to feature extraction from segmented images where we will propose approaches for the classification task.

Acknowledgements. This research was supported by the CONACYT Mastering Scholarship 613736. We thank to the anonymous reviewers for their contribution to improve this work.

References

1. World Health Organization: Cancer: fact sheet N°297 (2018). <http://www.who.int/mediacentre/factsheets/fs297/en/>
2. Lehman, C.D., Schnall, M.D.: Imaging in breast cancer: magnetic resonance imaging. *Breast Cancer Research*. 7(5), 215–219 (2005)
3. Berg, W.A., Gutierrez, L., Ness Aiver, M.S., Carter, W.B., Bhargavan, M., Lewis, R.S., Ioffe, O.B.: Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology*. 233(3), 830–849 (2004)
4. Kuhl, C.K., Schrading, S., Leutner, C.C., Morakkabati-Spitz, N., Wardelmann, E., Fimmers, R., Schild, H.H.: Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *Journal of clinical oncology*. 23(33), 8469–8476 (2005)
5. United States Environmental Protection Agency: Radiation Health Effects (2018). <https://www.epa.gov/radiation/radiation-health-effects>
6. Usuki, H., Ikeda, T., Igarashi, Y., Takahashi, I., Fukami, A., Yokoe, T., Asaishi, K.: What Kinds Of Non-palpable Breast Cancer Can Be Detected By Thermography?. *Biomedical thermology: the journal of the Japanese Society of Thermorogy*. 18(4), 8–12 (1998)
7. Head, J.F., Elliott, R.L.: Infrared imaging: making progress in fulfilling its medical promise. *IEEE Engineering in Medicine and biology Magazine*. 21(6), 80–85 (2002)
8. Tejerina, A.: Aula de Habilidades y Simulacion en Patologia de la Mama. In: *Ademas Comunicacion Grafica*, Madrid, Spain, pp. 52–56 (2009)

9. Borchardt, T.B., Concia, A., Limab, R.C.F., Resminia, R., Sanchez, A.: Breast thermography from an image processing viewpoint: a survey. *Sig. Process.* 93(10), 2785–2803 (2013)
10. Omranipour, R., Kazemian, A., Alipour, S., Najafi, M., Alidoosti, M., Navid, M., Izadi, S.: Comparison of the Accuracy of Thermography and Mammography in the Detection of Breast Cancer. *Breast Care.* 11(4), 260–264 (2016)
11. Pavithra, P.R., Ravichandran, K.S., Sekar, K.R., Manikandan, R.: The Effect of Thermography on Breast Cancer Detection. *Systematic Reviews in Pharmacy.* 9(1), 10–16 (2018)
12. Marques, R.S., Conci, A., Perez, M.G., Andaluz, V.H., Mejia, T.M.: An approach for automatic segmentation of thermal imaging in Computer Aided Diagnosis. *IEEE Latin America Transactions.* 14(4), 1856–1865 (2016)
13. Villalobos-Montiel, A.J., Chacon-Murguia, M.I., Calderon-Contreras, J.D., Ortega-Maynez, L.: Automatic Segmentation of Regions of Interest in Breast Thermographic Images. In: *Mexican Conference on Pattern Recognition*, pp. 135–144. Springer, Cham (2015)
14. Silva, L.F., Saade, D.C.M., Sequeiros-Olivera, G.O., Silva, A.C., Paiva, A.C., Bravo, R.S., Conci, A.: A new database for breast research with infrared image. *J. Med. Imaging Health Inf.* 4(1), 92–100 (2014)
15. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics.* 9(1), 62–66 (1979)
16. Burger, W., Burge, M.J.: *Principles of digital image processing: fundamental techniques.* Springer Science & Business Media. (2010)
17. H. Usuki, Y. Onoda, S. Kawasaki, T. Misumi, M. Murakami, S. Komatsubara, S. Teramoto.: Relationship between thermographic observations of breast tumors and the DNA indices obtained by flow cytometry. *Biomedical Thermology.* 10 (4), 282–285 (1990)
18. Duarte, A., Carrão, L., Espanha, M., Viana, T., Freitas, D., Bárto, P., Almeida, H.A.: Segmentation algorithms for thermal images. *Procedia Technology.* 16, 1560–1569 (2014)
19. Hinojosa, S., Pajares, G., Cuevas, E., Ortega-Sanchez, N.: Thermal Image Segmentation Using Evolutionary Computation Techniques. In: *Advances in Soft Computing and Machine Learning in Image Processing*, pp. 63–88. Springer, Cham (2018)
20. Ali, M.A., Sayed, G.I., Gaber, T., Hassanien, A.E., Snasel, V., Silva, L.F.: Detection of breast abnormalities of thermograms based on a new segmentation method. In: *Computer Science and Information Systems (FedCSIS), 2015 Federated Conference*, pp. 255–261. IEEE. (2015)
21. Sathish D.: Asymmetry analysis of breast thermograms using automated segmentation and texture features. *Signal, Image and Video Processing.* 11 (4), 745–752 (2017)
22. Hankare, P., Shah, K., Nair, D., Nair, D.: Breast cancer detection using thermography. *Int. Res. J. Eng. Technol.* 4(3), 2395–2356 (2016)
23. Prakash, R.M., Bhuvaneshwari, K., Divya, M., Sri, K.J., Begum, A.S.: Segmentation of thermal infrared breast images using K-means, FCM and EM algorithms for breast cancer detection. In: *Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017 International Conference*, pp. 1–4. IEEE. (2017)
24. Sayed, G.I., Soliman, M., Hassanien, A.E.: Bio-inspired swarm techniques for thermogram breast cancer detection. In: *Medical Imaging in Clinical Applications*, pp. 487–506. Springer International Publishing. (2016)
25. Mejía, T.M., Pérez, M.G., Andaluz, V.H., Conci, A.: Automatic Segmentation and Analysis of Thermograms Using Texture Descriptors for Breast Cancer Detection. In: *Computer Aided System Engineering (APCASE), 2015 Asia-Pacific Conference*, pp. 24–29. IEEE. (2015)

Daniel Sánchez-Ruiz, Iván Olmos-Pineda, J. Arturo Olvera-López

26. Garduño-Ramón, M.A., Vega-Mancilla, S.G., Morales-Henández, L.A., Osornio-Rios, R.A.: Supportive Noninvasive Tool for the Diagnosis of Breast Cancer Using a Thermographic Camera as Sensor. *Sensors*. 17(3), 497–518 (2017)
27. Oliphant, T.E.: Python for Scientific Computing. *Computing in Science & Engineering*. 9(3), 10–20 (2007)
28. Oliphant, T.E.: A guide to NumPy, USA: Trelgol Publishing, (2006)
29. Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J.M., Boulogne, F., Warner, J.D., Yager, N., Yu, T.: scikit-image: image processing in Python. *PeerJ*, 2, 1–18 (2014)

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
Noviembre de 2018
Printing 500 / Edición 500 ejemplares

