Advances in Computing Science

Research in Computing Science

Series Editorial Board

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 80,** noviembre 2014. 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 80**, November 2014. 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.

Volume 80

Advances in Computing Science

Evelio Martínez Martínez,
José Ángel González Fraga,
Omar Álvarez Xochihua,
Juan Ivan Nieto Hipólito (eds.)

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2014

ISSN: 1870-4069

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

Indexed in LATINDEX and Periodica / Indexada en LATINDEX y Periódica

Printing: 500 / Tiraje: 500

Printed in Mexico / Impreso en México

Preface

This Volume contains 10 extended versions of the selected articles presented at the 7th International Conference in Computer Science (CiComp, 2014), addressing research trends and advances in the area of Computer Science.

Review Committee members, based on a double-blind process, have carefully selected the papers contained in this Volume. The main criteria for selection were originality, technical quality and contribution to the field of Computer Science. Each paper was reviewed by three members of the reviewing committee.

The papers in this issue of the journal Research in Computing Science can be of interest to researchers and students in Computer Science, mainly for those interested in its application in everyday life, as well as for those that would like to know more about this fascinating themes.

This Volume is the result of the hard work of many people. In the first place, we would like to thank the contributors for the technical excellence of their papers that ensure the high quality of this publication. We would also like to extend a special thank to the members of the international editorial board and the additional reviewers for their insightful comments on the papers, which helped to improve the quality of this work

The submission, review and selection process was conducted on the basis of the free system OpenConf, www.openconf.com.

Evelio Martínez Martínez,
José Ángel González Fraga,
Omar Alvarez Xochihua,
Juan Ivan Nieto Hipólito
November 2014

Table of Contents

Page

A Novel Fourth Heart Sounds Segmentation Algorithm Based on Matching Pursuit and Gabor Dictionaries

Carlos Nieblas[1], Roilhi Ibarra[2] and Miguel Alonso[2]

[1] Meltsan Solutions,
Blvd. Adolfo Lopez Mateos No. 36 Segundo Piso Col. Alfonso XIII Mexico D.F.
[2] Departamento de Electrónica y Telecomunicaciones, División de Física Aplicada,
Centro de Investigación Científica y de Educación Superior de Ensenada,
Carretera Ensenada-Tijuana No. 3918, Ensenada, B.C., Mexico.
carlos.nieblas@meltsan.com, roilhi@cicese.edu.mx, aalonso@cicese.edu.mx

Abstract. In this paper we propose an efficient method for S4 heart sound segmentation based on the Matching Pursuit algorithm and Gabor Dicitonaries. An evaluation of this algorithm, through the use of different cardiac cycle events for S4 heart sound signals, showed a high performance, achieving a detection rate of 100% with the use of Gabor dictionary. The proposed method is practical, thus making it suitable for several applications such as signal compression and audio coding.
Keywords: Gabor, Matching Pursuit, Segmentation

1 Introduction

According to the World Health Organization (WHO) heart diseases represent the number one cause of death in the world: more people die every year from heart diseases than from any other cause [1]. Heart sound is a key signal to assess the mechanical functionality state of the heart. Auscultation is one of the simplest, quickest and, specifically, low cost effective techniques to analyze heart sounds. It is used to identify and diagnose a large number of heart pathologies simply by listening to cardiac sounds called "phonocardiograms" [2].

Thanks to recent advancements in data recording technology and digital signal processing it is now possible to record and analyze sound signals emitted by the heart. However, a computer analysis of the acoustic signals of the heart requires the different components of the heart cycle to be divided and timed separately. This is why we have developed a computerized method that allows such segmentation of both normal and abnormal heart sounds, first (S1) and second (S2), respectively. The analysis of systolc and diastolic murmurs is based on the frequency and location of the signal to determine the characteristics of the heart sounds in the time-frequency levels [3]. Methods based on Wavelet's theory have been used in the past to decompose, analyze and reconstruct heart sound signals [4]. Nevertheless, the robustness of reconstruction and compression

in decomposition that Matching Pursuit (MP) offers has provided with a powerful tool to represent a phonocardiogram through a linear combination of well concentrated waveforms, also called atoms, while at the same time performing an extraction of components S1 and S2, called segmentation [5].

The purpose of the present study is to select the optimal atoms to be used in thereconstruction of a first heart sound (S1) and to propose a segmentation algorithm of extra sounds near S1 using the matching pursuit algorithm and Gabor atoms.

2 Matching Pursuit Algorithm

Matching Pursuit (MP) is an iterative greedy algorithm that decomposes a signal $x(t)$ into a sparse linear combination of waveforms that belong to a redundant dictionary of functions called time-frequency atoms and residual term $R_M(t)$ [6,7]. MP aims at to find sparse decompositions of signals, that is, to obtain a representation that accounts for most or all information of a signal with a linear combination of a small number of elementary waveforms (called atoms). The decomposition is performed by projecting the signal $x(t)$ over a redundant dictionary of functions $D = g_\gamma(t), \gamma \epsilon \Gamma$ where $g_\gamma(t), \gamma \epsilon \Gamma^2(R), \Gamma = R^+ \times R^2$ and by selecting the atoms which can best match the local structure of the signal. The signal $x(t)$ can be reconstructed from the sum of M atoms and a residual term $R_M(t)$ as

$$x(t) = \sum_{m=1}^{M} \alpha_m \cdot g_{\gamma_m}(t) + R_M(t) \tag{1}$$

where $g_(\gamma_m)$ and α_m are the $m-th$ optimal and ponderation factor respectively. MP is an iterative descent algorithm which selects the optimal atom at each iteration [8]., see Algorithm 1.

Algorithm 1 Standard Matching Pursuit

input: $x(t)$; $\mathcal{D} = \{g_\gamma(t), \gamma \in \mathbf{\Gamma}\}$
output: $\alpha_m, g_{\gamma_m}(t)$

$R = x(t)$
$\alpha_m = 0$
for each m in M
$g_{\gamma m} = \arg\max_{\gamma \in \Gamma} |\langle R, g_\gamma \rangle|$
$\alpha_m = \langle R, g_{\gamma_m} \rangle$
$R = R - \alpha_m \cdot g_{\gamma_m}$
until target signal-to-noise ratio (SNR) or the M iteration has been reached.

3 Dictionary Selection

Selecting an adequate dictionary plays a fundamental role in the performance of the MP algorithm. Several kinds of time-frequency dictionaries have been proposed in literature, for instance: The Wavelet Packet, The Cosine packet, MDCT (Modified Discrete Cosine Transform), and Gabor dictionaries [8]. After conducting preliminary tests using all of the above mentioned dictionaries we have found the Gabor's functions can model the non-stationary heart sound signal. In addition to this, the ability of its functions to model heart sounds has been previously reported [8]. Gabor atoms are obtained by dilating, translating and modulating a mother window $w(t)$ which is generally revalued, positive and of unit norm $\int w(t)^2 dt = 1$.

$$g_\gamma(t) = \frac{1}{\sqrt{s}} w(\frac{t-u}{s}) e^{i2\pi\xi(t-u)} \tag{2}$$

where w is the Gaussian window $w(t) = \sqrt[4]{2}e^{-\pi t^2}$, the scale s is used to control the width of the waveform envelope, the time displacement u is used to specify the temporal location of the atom and ξ is the modulating frequency. We define $\gamma_m = (s_m, u_m, \xi_m)$, where the index γ_m is an element of the set $\mathbf{\Gamma}$.

3.1 Segmentation Algorithm

We approximate the optimal attacks using the matching pursuit algorithm and Gabor dictionaries. Then, we look for all the x_j optimal attacks that satisfy the condition

$$\mathcal{O}(x_j) > \mu$$

where μ is the duration of S2 heart sounds i.e. = 0.12s. The optimal attacks were selected using the heart sound segmentation algorithm based on matching pursuit [5].

Once MP decomposition of heart sound signal has been performed, it produces a set of Gabor atoms. These Gabor atoms are optimal to represent the analyzed signal and are then stored in a so-called book. Te first step of the segmentation algorithm consists of selecting the position u_m for all optimal atoms.

The next step consists of identifying the minimum position for the analyzed signal. The Gabor atom with the minimum position is considered to be an onset and it is stored into θ_j. Fig. 1 shows the onset detected between S1 and S4 heart sounds.

The whole segmentation process is described in the form of pesudo code in Algorithm 2.

Algorithm 2 Segmentation

input: $book = \{g_{\gamma_m}, \gamma \in \mathbf{\Gamma}\}$
ouput: $o_i = 0, \theta_i = 0$
repeat j
x_j
repeat n=1:3
$o_i = min(u_n)$
end
$\theta_j = o_i$
end

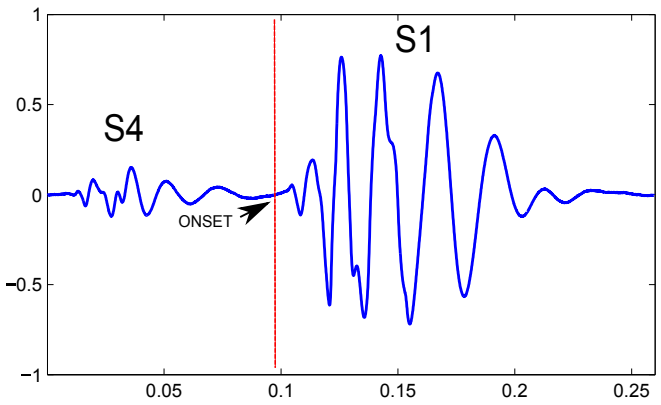


Fig. 1. Onset detection of S1 in a PCG waveform using the algorithm. It is shown also the waveform of S4 before this.

3.2 Test Corpus

The heart sound signals used in this study were all obtained from Littmann® website. Seven cardiac cycle events were analyzed, containing normal sounds (normal split S1, normal split S2) and s4 sounds. Five dictionaries were implemented for the S4 heart sounds segmentation. Table 1 shows a list of the dictionaries used.

A total of 240 simulations were carried out using Gabor dictionaries as well as four different iteration values. All onsets of the heart sounds were registered by hand separately by two different people, then results were averaged. The annotation process consisted of repetitively listening to the sound signal using headphones while visually inspecting the time waveform.

Table 1. Multi-block Gabor dictionaries implemented for segmentation of S4 heart sounds.

Dictionary	window size (%)	Size of atom	total of atoms
1	50	128,256,512	23639
2	50	64,128,256,512,1024	39038
3	50	32,64,128	24760
4	50	256,512,1024	22761
5	50	1024	7182

Table 2. Number of onsets detected using Gabor multi-block dictionaries.

onset identified	onset detected	Number of iteration	Gabor dictionary	Total of atoms
7	4	2	1	23639
7	5	2	2	39038
7	0	2	3	24760
7	5	2	4	22761
7	5	2	5	7182
7	7	3	1	23639
7	4	3	2	39038
7	0	3	3	24760
7	4	3	4	22761
7	3	3	5	7182
7	5	4	1	23639
7	4	4	2	39038
7	0	4	3	24760
7	4	4	4	22761
7	3	4	5	7182
7	3	5	1	23639
7	2	5	2	39038
7	0	5	3	24760
7	2	5	4	22761
7	3	5	5	7182

3.3 Evaluation

All the simulations carried out in the present work were performed under Matlab©. Heart sounds were decomposed using MPTK (Matching Pursuit ToolKit)[9]. MPTK is an open source package that provides a fast implementation of the Matching Pursuit algorithm.

We present a number of simulation experiments to assess the performance of the proposed algorithm for heart sound segmentation. The tolerance criterion between the start of the attack and the detected onset was set to 0.01 s. Table 2 shows the results obtained from each cardiac cycle event as well as the number of atoms in each dictionary and detection of onsets. Although the dataset used to evaluate the algorithm is small, we consider the results highly encouraging since it exhibits high performance for onset detection. During our simulation we found that by using Gabor dictionaries with lengths of atoms [128,256, 512] and three optimal atoms 100% of the onsets can be detected. Three Gabor atoms were required to obtain 75% of the reconstruction of S1 heart sound. Using three Gabor atoms the residual energy of S1 heart sound is less than the energy of S4 heart sound. Then we can reconstruct only the S1 heart sounds. Fig 2. shows the S4 and S1 heart sounds segmentation using the dictionary parameters above mentioned. The time-frecuency representation of S1 heart sounds using three atoms is showed in Fig. 3.

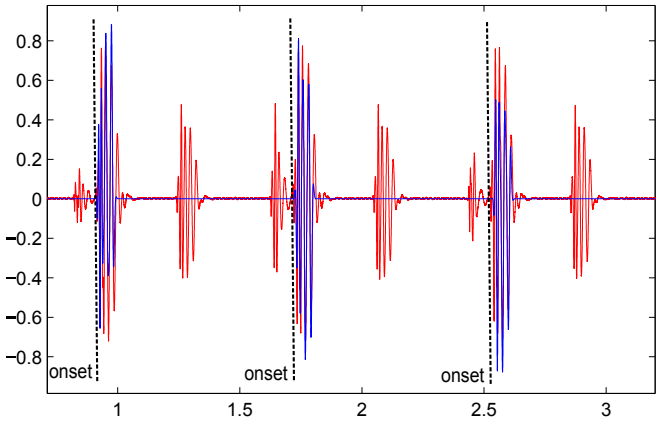


Fig. 2. Section of a S4l heart sound illustrating the main cardiac cycle events, namely the first heart sound (SI), the systolic period, the second heart sound (S2), the diastolic period.

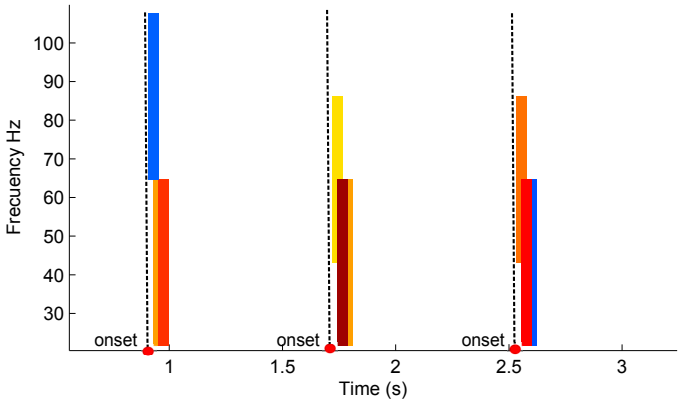


Fig. 3. Time-frequency representation of a normal heart sound obtained with the MP algorithm. Each of the rectangles (also called boxes) in this representation is a Gabor atom.

4 Conclusions

In this paper, we presented an efficient S4 heart sound segmentation method based on the Matching Pursuit and Gabor dictionaries. We have shown that Gabor dictionaries provide an efficient segmentation of S4 heart sound signals. The technique showed high performance achieving a detection rate of 100% for onsets between S1 and S4. The method presented can be directly used in the pre-processing step of PCG signal compression, and PCG based, audio coding, source separation and PCG signal enhancement.

Acknowledgment. We thank Meltsan Solutions and CICESE for their help through the annotation and validation process. The authors also wish to thank the team that develops MPTK for making their algorithms available.

References

1. Organization, W.H.: Global status report on noncommunicable diseases 2010 (Geneva, Switzerland 2011)
2. Abbas, A., Bassam, R.: Phonocardiography Signal Processing. Synthesis Lectures on Biomedical Engineering Series. Morgan & Claypool (2009)
3. Boutana, D., Benidir, M., Barkat, B.: Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis. IET signal processing **5** (2011) 527–537
4. Martinez-Alajarin, J., Ruiz-Merino, R.: Efficient method for events detection in phonocardiographic signals. In: Microtechnologies for the New Millennium 2005, International Society for Optics and Photonics (2005) 398–409

5. Nieblas C.I., Alonso M.A., C.R.V.S.: High performance heart sound segmentation algorithm based on matching pursuit. IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE). (2013) 96–100
6. Mallat S., Z.Z.: Matching pursuit with time-frequency dictionaries. IEEE Transactions on signal processing **41** (1993) 3397–3415
7. R., G., E., B.: Harmonic decomposition of audio signals with matching pursuit. IEEE Transactions on processing **51** (2003) 101–111
8. Ravelli E., R.G., L., D.: Union of mdct bases for audio coding. IEEE Transactions on audio, speech, and language processing **16** (2008) 1364–1372
9. Krstulovic, S., Gribonval, R.: Mptk: Matching pursuit made tractable. In: Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on. Volume 3. (2007) III–III

Simulator Based on a Simple Biped System

Sandra Cuatlaxahue Formacio and Pablo Sánchez-Sánchez

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Electrónica
robotics group oocelo
Puebla, México
cformaciosd@gmail.com
lepable@ece.buap.mx

Abstract. This study shows the mathematical modeling and the development of the simulator of a simple biped system in **Matlab**©. We used specific libraries of **Matlab**© that allowed us to simulate mechanical systems. In order to design the 3D model, we used **SolidWorks**©. The biped system is based on the structure of lower limb exoskeleton which is used in medical rehabilitation. We present the dynamic model calculation of the biped system through Euler-Lagrange method, and the stability analysis using Lyapunov theory. We present the implementation of a tracking control structure using a trajectory defined by fifth-order polynomials. The main consideration in this work is that the system is free of interaction with the environment, i.e. , we discussed the ideal case.

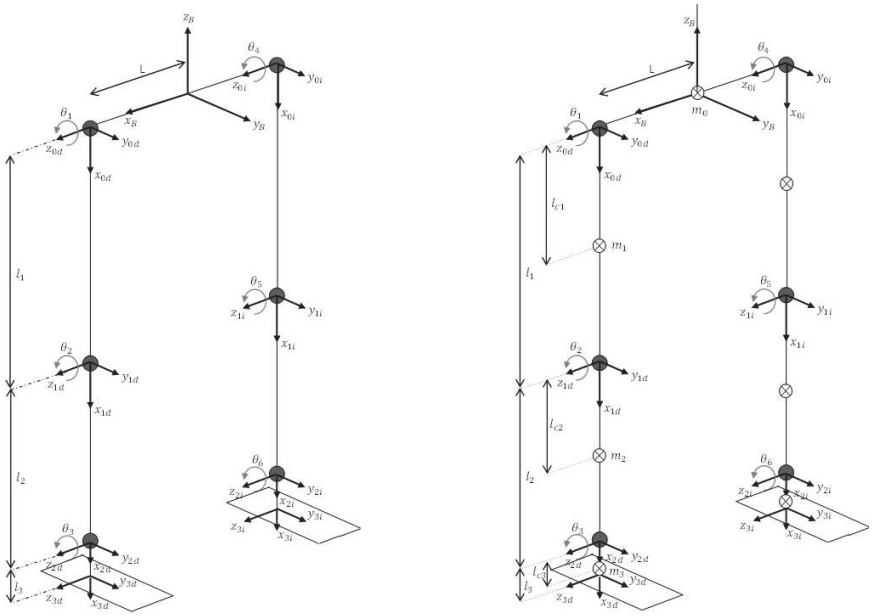
Keywords: Simulator, dynamic model, biped system, lower limbs, fifth-order polynomial

1 Introduction

A computer simulation is an attempt to model a real-life or hypothetical situation on a computer so that it can be studied to see how the system works. By changing variables in the simulation, predictions may be made about the behaviour of the system. Simulator is a program that simulates specific conditions or the characteristics of a real process or machine for the purposes of research or study. Currently, simulators are being increasingly used in different areas of knowledge. Simulators in engineering are used in the design of complex systems. The results obtained in the analysis of data allow us to understand and improve the performance of the system[1]. An approach in robotics is the autonomous or semi-autonomous system design with ability to interact with its environment. Since complexity of robot control is increasing, simulators are becoming essential tools to understand the behavior of the system. The simulation allow us to enhance the design of the system and eliminate mechanical failures, before building the prototype[2, 3]. The study of the bipedal robots has been around for more than three decades and there are still problems to be solved. Accordingly, this paper is focused on the development of a simulator based on a simple biped system with the purpose to study the major physical characteristics.

2 Modeling of the System

The fixed frame is located at the base of the hip. The system is divided into 2 subsystems, right leg $\{SPD\}$ and left leg $\{SPI\}$. The first step is to obtain the forward kinematics of the system. The forward kinematics consists in obtaining the spatial location of the links with respect to a fixed coordinate system [4]. In order to analyze the system we requires a motion diagram, **Fig. 1(a)**. The diagram is used to obtain the forward kinematics.



(a) ***coordinate system*** **(b)** ***Mass distribution***

Fig. 1. Biped system frame

The forward kinematics is the first and most important step for the application of Euler-Lagrange method. The diagram of mass distribution is used for the analysis of velocity for each leg, **Fig. 1(b)**.

2.1 Forward Kinematics Model

In order to obtain the homogeneous transformation matrices (spatial location of each link) we used the *Denavit-Hartenberg algorithm* [5]. The location of each link is obtained with respect to the fixed reference system located at the base of the system, **Fig. 1(a)**.

The transformation matrices for $\{SPD\}$ are the following:

$$ {}^{B}A_{0d} = \begin{bmatrix} 0 & 0 & 1 & L \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1} $$

$$ {}^{B}A_{1d} = \begin{bmatrix} 0 & 0 & 1 & L \\ \sin(\theta_1) & \cos(\theta_1) & 0 & l_1 \sin(\theta_1) \\ -\cos(\theta_1) & \sin(\theta_1) & 0 & -l_1 \cos(\theta_1) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2} $$

$$ {}^{B}A_{2d} = \begin{bmatrix} 0 & 0 & 1 & L \\ S\theta_{12} & C\theta_{12} & 0 & l_2 S\theta_{12} + l_1 S\theta_1 \\ -C\theta_{12} & S\theta_{12} & 0 & -l_2 C\theta_{12} - l_1 C\theta_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3} $$

$$ {}^{B}A_{3d} = \begin{bmatrix} 0 & 0 & 1 & L \\ S\theta_{123} & C\theta_{123} & 0 & l_3 S\theta_{123} + l_2 S\theta_{12} + l_1 S\theta_1 \\ -C\theta_{123} & S\theta_{123} & 0 & -l_3 C\theta_{123} - l_2 C\theta_{12} - l_1 C\theta_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4} $$

where $S\theta_1 = \sin(\theta_1)$, $C\theta_1 = \cos(\theta_1)$, $S\theta_{12} = \sin(\theta_1 + \theta_2)$, $C\theta_{12} = \cos(\theta_1 + \theta_2)$, $C\theta_{123} = \cos(\theta_1 + \theta_2 + \theta_3)$ y $S\theta_{123} = \sin(\theta_1 + \theta_2 + \theta_3)$.
$\{SPI\}$ has the same equations that $\{SPD\}$, the only difference is the position of the x-axis since L is now $-L$ and it is evaluated with the corresponding angles, as seen in **Fig. 1(a)**.

2.2 Dynamic Model

In the design of robots, control algorithms and simulators is important to consider the equations of motion [5]. We introduce the Euler-Lagrange equations, which describe the evolution of a mechanical system. The equation of motion is the following:

$$ \frac{d}{dt}\left[\frac{\partial \mathcal{L}\left(\theta, \dot{\theta}\right)}{\partial \dot{\theta}}\right] - \frac{\partial \mathcal{L}\left(\theta, \dot{\theta}\right)}{\partial \theta} = \tau \tag{5} $$

where $\theta \in \mathbb{R}^n$ is the vector of generalized joint coordinates, $\tau \in \mathbb{R}^n$ is the vector of torques that act in the joints, and $\mathcal{L}\left(\theta, \dot{\theta}\right)$ is know as Lagrangian, which is the difference between the kinetic and potential energy:

$$ \mathcal{L}\left(\theta, \dot{\theta}\right) = \mathcal{K}\left(\theta, \dot{\theta}\right) - \mathcal{U}(\theta) \tag{6} $$

Considering that the biped system had no interaction with the environment, we can identify the input variables as the applied torques and the output variables as the positions[5].

By solving the Euler-Lagrange equation of motion described in (5) we obtain the dynamic model. The dynamic model is defined as:

$$M(\theta)\ddot{\theta} + C\left(\theta, \dot{\theta}\right)\dot{\theta} + g(\theta) = \tau \tag{7}$$

where $\theta \in \mathbb{R}^n$ is the vector of generalized joint coordinates, $M(\theta) \in \mathbb{R}^{n \times n}$ is the symmetric positive definite inertia matrix, $C\left(\theta, \dot{\theta}\right) \in \mathbb{R}^{n \times n}$ is the matrix of centripetal and Coriolis torques, $g(\theta) \in \mathbb{R}^n$ is the vector of gravitational torques, $\tau \in \mathbb{R}^n$ is the vector of torques that act in the joints. Assuming for simplicity that all robots have revolute joints, we can set the following properties [5]:

Property 1 *$M(\theta)$ satisfies $\lambda_m \|x\|^2 \leq x^T M x \leq \lambda_M \|x\|^2 \; \forall \theta$, $x \in \mathbb{R}^{n_i}$ where $\lambda_m \triangleq \min_{\forall \theta \in \mathbb{R}^{n_i}} \lambda_{\min}(M)$, $\lambda_M \triangleq \max_{\forall \theta \in \mathbb{R}^{n_i}} \lambda_{\max}(M)$, and $0 < \lambda_m \leq \lambda_M < \infty$.*

△

Property 2 *$M(\theta)$ satisfies $0 < \lambda_m \leq \|M(\theta)\| \leq \lambda_M < \infty$.*

△

Property 3 *M^{-1} satisfies $\sigma_m \|x\|^2 \leq x^T M^{-1} x \leq \sigma_M \|x\|^2 \forall \theta, x \in \mathbb{R}^{n_i}$, $\sigma_m \triangleq \min_{\forall \theta \in \mathbb{R}^{n_i}} \lambda_{\min}\left(M^{-1}\right)$, and $0 < \sigma_m \leq \sigma_M < \infty$. M^{-1} satisfies $0 < \sigma_m \leq \left\|M^{-1}\right\| \leq \sigma_M < \infty$.*

△

Property 4 *$M(\theta)$ has the following relationship with the kinematic energy $\mathcal{K}\left(\theta, \dot{\theta}\right) = \frac{1}{2}\dot{\theta}^T M(\theta)\dot{\theta}$.*

△

Property 5 *The vector $C(\theta, x)\, y$ satisfies $C(\theta, x)\, y = C(\theta, y)\, x \; \forall x, y \in \mathbb{R}^n$.*

△

Property 6 *With the proper definition of $C\left(\theta, \dot{\theta}\right)$, the matrix $\frac{1}{2}\dot{\theta}^T\left[\dot{M}(\theta) - 2C\left(\theta, \dot{\theta}\right)\right]\dot{\theta} \equiv 0$ is skew-symmetric.*

△

Property 7 *$C^T\left(\theta, \dot{\theta}\right) = \frac{1}{2}\frac{\partial}{\partial \theta}\left[\dot{\theta}^T M(\theta)\dot{\theta}\right]$.*

△

The dynamic model for $\{SPD\}$ is defined as:

$$\begin{bmatrix} m_{11}\ m_{12}\ m_{13} \\ m_{12}\ m_{22}\ m_{23} \\ m_{13}\ m_{23}\ m_{33} \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \end{bmatrix} + \begin{bmatrix} c_{11}\ c_{12}\ c_{13} \\ c_{21}\ c_{22}\ c_{23} \\ c_{31}\ c_{32}\ c_{33} \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} \tag{8}$$

where

$$\begin{aligned} m_{11} &= m_1 l_{c1}^2 + m_2 l_1^2 + m_3 l_1^2 + m_2 l_{c2}^2 + m_3 l_2^2 + 2\left(m_2 l_{c2} l_1 + m_3 l_2 l_1\right)\cos\left(\theta_2\right) \\ &\quad + m_3 l_{c3}^2 + 2 m_3 l_{c3} l_2 \cos\left(\theta_3\right) + 2 m_3 l_{c2} l_1 \cos\left(\theta_2 + \theta_3\right) + I_1 + I_2 + I_3 \quad (9) \\ m_{12} &= m_2 l_{c2} + m_3 l_2^2 + \left(m_2 l_{c2} l_1 + m_3 l_2 l_1\right)\cos\left(\theta_2\right) + 2 m_3 l_{c3} l_2 \cos\left(\theta_3\right) \\ &\quad + m_3 l_{c3} l_1 \cos\left(\theta_2 + \theta_3\right) + m_3 l_{c3}^2 + I_2 + I_3 \quad (10) \\ m_{13} &= m_3 l_{c3}^2 + m_3 l_{c3} l_2 \cos\left(\theta_3\right) + m_3 l_{c3} l_1 \cos\left(\theta_2 + \theta_3\right) + I_3 \quad (11) \\ m_{22} &= m_2 l_{c2}^2 + m_3 l_2^2 + m_3 l_{c3}^2 + 2 m_3 l_{c3} l_2 \cos\left(\theta_3\right) + I_2 + I_3 \quad (12) \\ m_{23} &= m_3 l_{c3}^2 + m_3 l_{c3} l_2 \cos\left(\theta_3\right) + I_3 \quad (13) \\ m_{33} &= m_3 l_{c3}^2 + I_3 \quad (14) \end{aligned}$$

$$\begin{aligned} c_{11} &= -2\left(m_2 l_{c2} l_1 + m_3 l_2 l_1\right)\sin\left(\theta_2\right)\dot{\theta}_2 - 2 m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 \\ &\quad + 2 m_3 l_{c3} l_1 \sin\left(\theta_2 + \theta_3\right)\left(\dot{\theta}_2 + \dot{\theta}_3\right) \quad (15) \\ c_{12} &= -2 m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 - m_3 l_{c3} l_1 \sin\left(\theta_2 + \theta_3\right)\left(\dot{\theta}_2 + \dot{\theta}_3\right) \quad (16) \\ c_{13} &= -m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 - m_3 l_{c3} l_1 \sin\left(\theta_2 + \theta_3\right)\left(\dot{\theta}_2 + \dot{\theta}_3\right) \quad (17) \\ c_{21} &= \left(m_2 l_{c2} + m_3 l_2\right) l_1 \sin\left(\theta_2\right)\dot{\theta}_1 + m_3 l_{c3} l_1 \sin\left(\theta_2 + \theta_3\right)\dot{\theta}_1 \\ &\quad - 2 m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 \quad (18) \\ c_{22} &= -2 m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 \quad (19) \\ c_{23} &= -m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_3 \quad (20) \\ c_{31} &= m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_1 + 2 m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_2 + m_3 l_{c3} l_1 \sin\left(\theta_2 + \theta_3\right)\dot{\theta}_1 \quad (21) \\ c_{32} &= m_3 l_{c3} l_2 \sin\left(\theta_3\right)\dot{\theta}_2 \quad (22) \\ c_{33} &= 0 \quad (23) \end{aligned}$$

$$\begin{aligned} g_1 &= \left[m_1 l_{c1} + m_2 l_1 + m_3 l_1\right]\sin\left(\theta_1\right) + \left[m_2 l_{c2} + m_3 l_2\right]\sin\left(\theta_1 + \theta_2\right) \\ &\quad + m_3 l_{c3} \sin\left(\theta_1 + \theta_2 + \theta_3\right) \quad (24) \\ g_2 &= \left[m_2 l_{c2} + m_3 l_2\right] g \sin\left(\theta_1 + \theta_2\right) + m_3 l_{c3} g \sin\left(\theta_1 + \theta_2 + \theta_3\right) \quad (25) \\ g_3 &= m_3 l_{c3} g \sin\left(\theta_1 + \theta_2 + \theta_3\right) \quad (26) \end{aligned}$$

Owing to similarity of the reference system definition for both subsystems, $\{SPI\}$ has the same equations as $\{SPD\}$, except that it was evaluated with θ_4, θ_5 and θ_6.

2.3 Control Structure

In order to move the joints from an initial point to desired point (position control) or move the joints along a defined trajectory (control of movement), we applied a control structure, which allows the system to perform the assigned task [5]. The control structure is defined as follows:

$$\tau = K_p\tilde{\theta} + K_v\dot{\tilde{\theta}} + g\left(\theta\right) \tag{27}$$

where $K_p, K_v \in \mathbb{R}^{n\times n}$ are the proportional and derivative gains, respectively. The control structure described in (27) use the positions, velocity and gravitational torque information [5]. Note that the system is controlled in closed-loop as shown in **Fig. 2**.

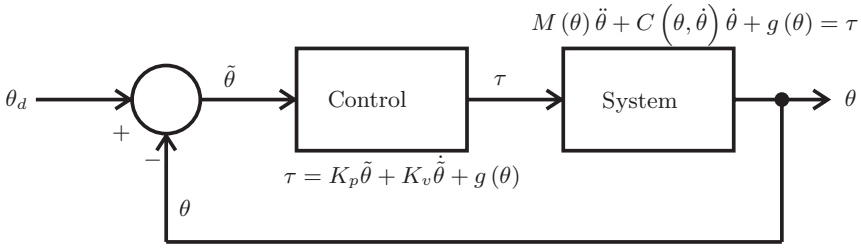


Fig. 2. System with control loop (inputs and outputs)

Demonstration of Stability. In order to demonstrate the stability of the system and the control structure, we use the Lyapunov theory. The first step is to design the matrices K_p and K_v such that the position error $\tilde{\theta}$ asymptotically vanishes, i. e. $\lim_{t\to\infty}\tilde{\theta}(t) = 0 \in \mathbb{R}^n$[5].
The closed-loop system equation obtained by combining the dynamic model described in (7), and control structure in (27), can be written as:

$$\frac{d}{dt}\begin{bmatrix}\tilde{\theta}\\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}-\dot{\theta}\\ M(\theta)^{-1}\left[K_p\tilde{\theta} - K_v\dot{\theta} - C(\theta,\dot{\theta})\dot{\theta}\right]\end{bmatrix} \tag{28}$$

(28) is known as *closed-loop equation*, which is an autonomous differential equation. Therefore, the origin of the state space is its unique equilibrium point[5]. In order to carry out the stability analysis of (28), the following Lyapunov function candidate based on the *energy shaping methodology* was proposed [5]:

$$V(\dot{\theta},\tilde{\theta}) = \frac{\dot{\theta}^T M(\theta)\dot{\theta}}{2} + \frac{\tilde{\theta}^T K_p\tilde{\theta}}{2} \tag{29}$$

Since $M(\theta)$ is a positive definite matrix the first term of $V(\dot{\theta},\tilde{\theta})$ is a positive definite function with respect to $\dot{\theta}$. The second one term is a positive definite

function with respect to position error $\tilde{\theta}$, because K_p is a positive definite matrix. Therefore $V(\dot{\theta}, \tilde{\theta})$ is a globally positive definite and radially unbounded function[5]. The derivative of Lyapunov function candidate (29) along the trajectories of the closed-loop (28) is:

$$\dot{V}(\dot{\theta}, \tilde{\theta}) = \dot{\theta}^T M(\theta)\ddot{\theta} + \frac{\dot{\theta}^T \dot{M}(\theta)\dot{\theta}}{2} + \tilde{\theta}^T K_p \dot{\tilde{\theta}} \tag{30}$$

and after some algebra and using the **Property 6** it can be written as:

$$\dot{V}(\dot{\theta}, \tilde{\theta}) = -\dot{\theta}^T K_v \dot{\theta} \leq 0, \tag{31}$$

which is a globally negative semidefinite function. Therefore, we concluded stability of the equilibrium point. In order to prove asymptotic stability, we applied the *LaSalle invariance principle*

$$\dot{V}(\dot{\theta}, \tilde{\theta}) < 0. \tag{32}$$

In the region

$$\Omega = \left\{ \begin{bmatrix} \tilde{\theta} \\ \dot{\theta} \end{bmatrix} \in \mathbb{R}^n : V(\tilde{\theta}, \dot{\theta}) = 0 \right\} \tag{33}$$

the unique invariant is $\begin{bmatrix} \tilde{\theta}^T & \dot{\theta}^T \end{bmatrix}^T = 0 \in \mathbb{R}^{2n}$[5].

2.4 Planning of Trajectory

The biped system consisting of hip, knee and ankle must perform the movements of the human walking. In order to define the desired trajectory, we used the sagittal-plane motion of the lower extremity, **Fig. 3** [6].

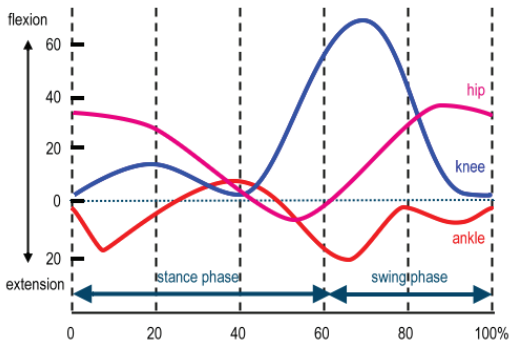


Fig. 3. Sagittal-plane motion of lower extremity during one gait cycle

The trajectory of human walking was define using fifth order polynomials. These polynomials restrict the positions, velocities and accelerations and generate soft movements [7]. A path from θ_{i} to θ_{f} is defined as a continuous map, $\tau : [0, 1] \rightarrow Q$, with $\tau(0) = \theta_{\mathrm{i}}$ and $\tau(1) = \theta_{\mathrm{f}}$. A trajectory is a function of time $\theta(t)$ such that $\theta(t_0) = \theta_{\mathrm{i}}$ and $\theta(t_f) = \theta_{\mathrm{f}}$. In this case, $t_f - t_0$ represents the amount of time taken to execute the trajectory. Since the trajectory is parameterized by time, we can compute velocities and accelerations along the trajectories by differentiation. If τ is time-dependent then a path is a special case of a trajectory, one that will be executed in one unit of time. In other words, in this case, τ gives a complete specification of the robot's trajectory, including the time derivatives (knowing the differentiate τ, we can obtain time derivatives) [7].

3 Design of the 3D System

The proposed biped system has 6 degrees of freedom. The joints of the system are rotational. In the middle of the hip is located the base of the system, **Fig. 4(a)**. The biped system is based on lower limb exoskeleton, **Fig. 4(b)**. The system is drawn in **SolidWorks©**.

(a) *SolidWorks© design* **(b) *Exoskeleton of lower limbs***

Fig. 4. Biped system

Prismatic joints (translation movement) is one of the options to design a *lower limb rehabilitation system*. The main advantage of this technique is the simplicity

in the modeling and implementation. However, designing a biped system considering rotational joints has the advantage of better leg kinematic description [8].
SolidWorks© allows us to applying the *finite element* analysis and define the physical properties of each element. These information is needed for movement modeling. **Matlab©** combines the physical properties, the dynamic model (mathematical representation of the system) and specific libraries in a programming environment [9].

4 Simulator Designing

We need several elements in order to develop the simulator: the *dynamic model* which describes the system's response to internal or external stimuli, the *control law* which controls the actions of the system to perform the task, the 3D model that allows us to visualize the system movements, and the physical parameters which give the system its physical characteristics, as seen in **Fig. 5**. The programming environment is **Matlab©**.

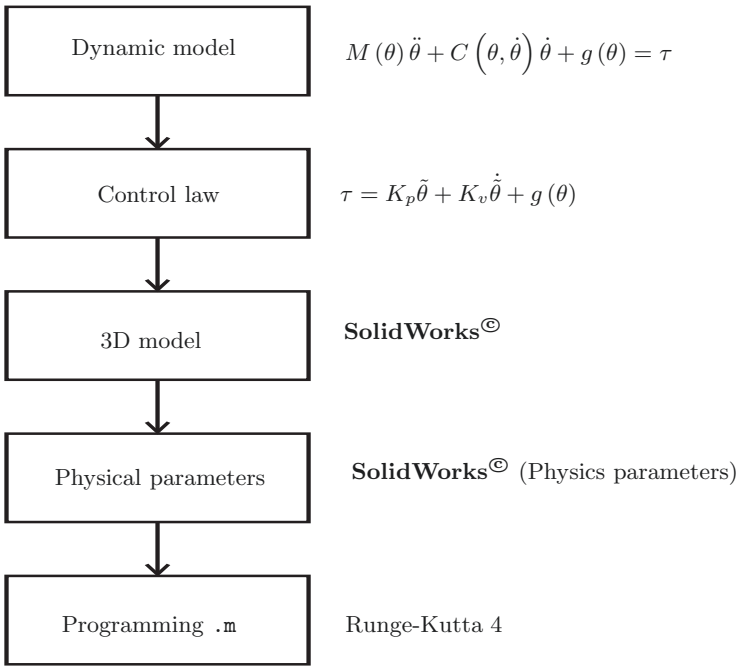


Fig. 5. Simulator elements

The programming language that set the numerical method is needed to relate these elements. **Matlab©**-Simulink uses Runge-Kutta 4 (`ode45`).

The simulator uses the following procedure:

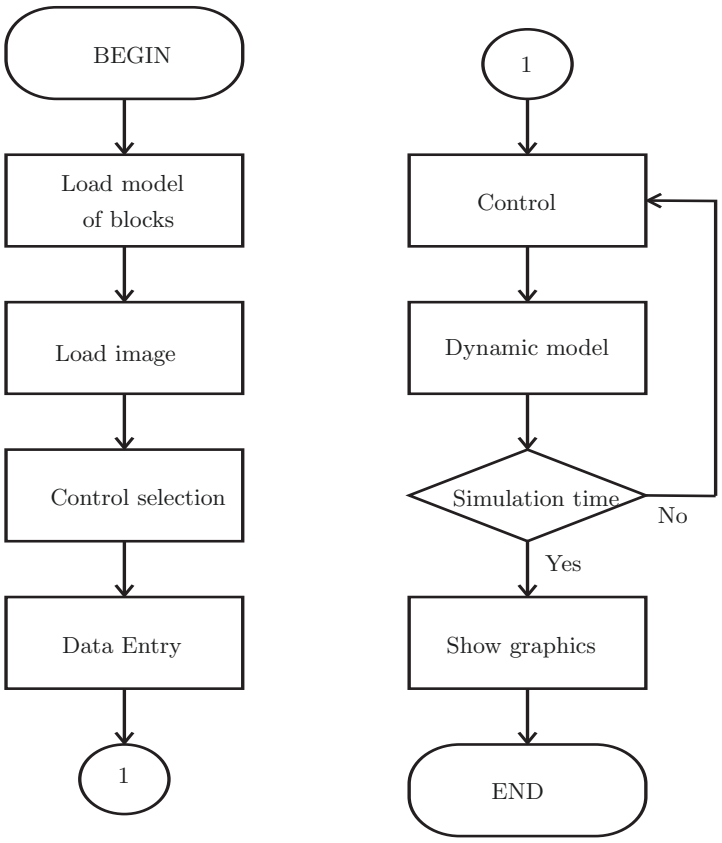


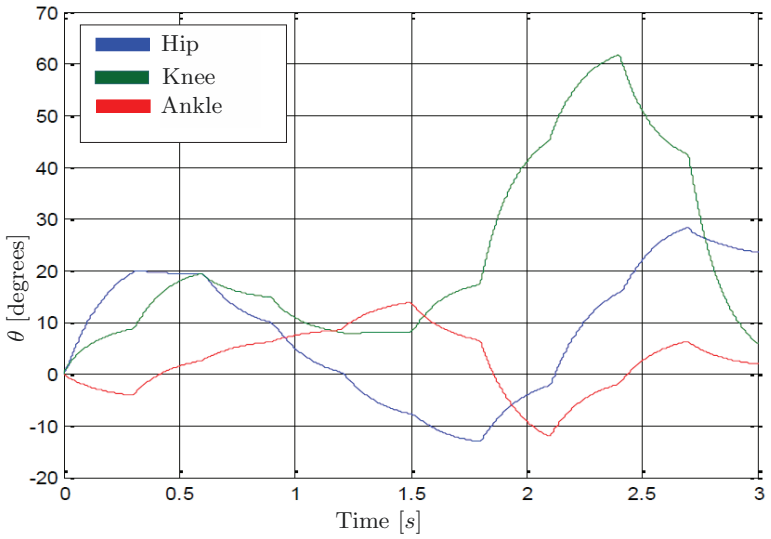
Fig. 6. Flow diagram

Matlab© allows us to use the 3D image made in **SolidWorks©** using the Toolbox call as *SimMechanics*. This toolbox converts the image into an object, this object can moves according to the results of the numerical method `ode45`.

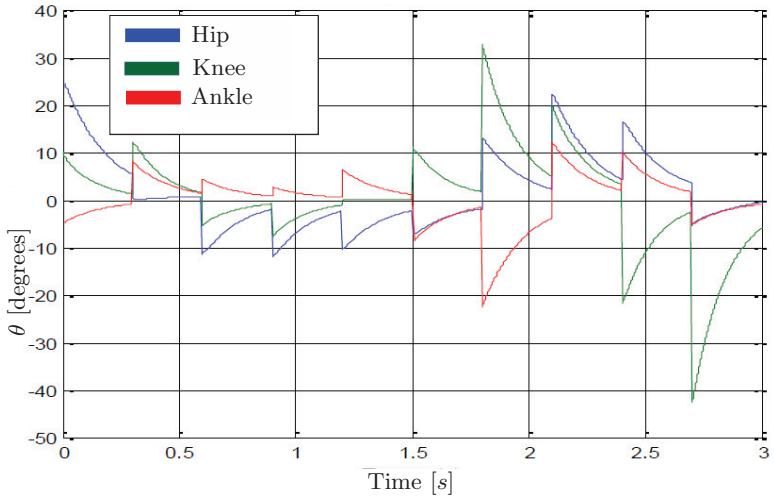
SimMechanics provides a multibody simulation environment for 3D mechanical systems. You model the multibody system using blocks representing bodies, joints, constraints, and force elements, and then *SimMechanics* formulates and solves the equations of motion for the complete mechanical system. Models from CAD systems requires mass, inertia, joint, constraint, and 3D geometry information. An automatically generated 3D animation lets us visualize the system dynamics.

5 Results

Positions for each joint of the right leg $\{SPD\}$ is shown in **Fig. 7**



(a) *Position*



(b) *Error position*

Fig. 7. Position and error position results (right leg $\{SPD\}$)

The trajectories obtained are similar to those shown in **Fig. 3**. To the left leg, the graphical output are similar but with a different start point to allow synchronization of the phases of support and oscillation of the movement of the legs in the running. **Fig. 8** shows the human gait shown by the simulator.

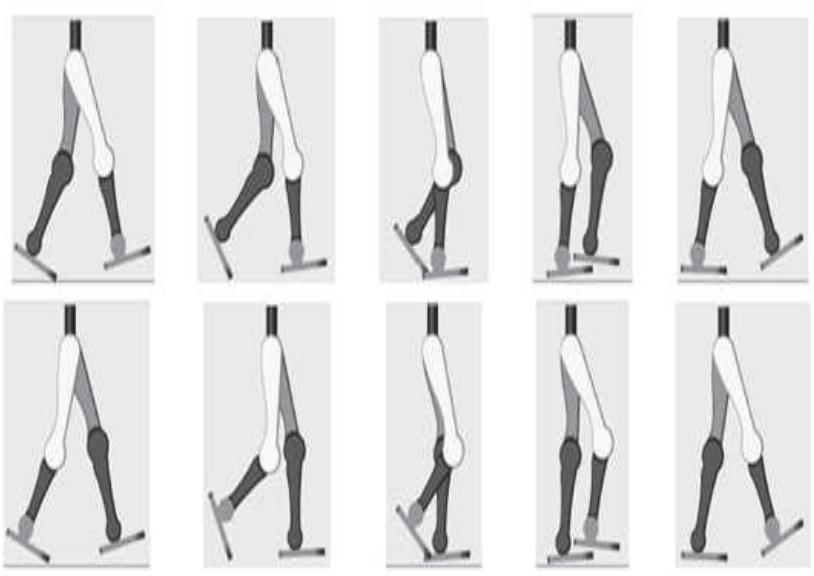


Fig. 8. Movement sequences

6 Conclusion

We developed a simulator using the system dynamic model, the system parameters, the implementation of a control law, the numerical method (Runge-Kutta), and the 3D drawing. The trajectory defined by the fifth-order polynomial adequately approaches the human gait. The trajectory can be improved by reducing the distance between the points that define it. The gain tuning of the control structure markedly contributes to the desired response and we can see its behavior as an amplifier with adjustable gain. The trajectory, defined by the fifth-order polynomial and the adequate selection of gains, allows the simulator efficiently replicate the human walking. Our results describe the system without interaction with the environment (contact with the floor, resistance to movement or other variables that depend on the time and velocity). The biped system analysis and the development of the simulator are the starting point for the development of a prototype with application in rehabilitation.

References

1. Moore, H.: MATLAB for Engineers. Prentice Hall (2014)
2. The MathWorks, `http://www.mathworks.es/discovery/robotica.html`
3. Kajita, S. and Espiau, B.: Legged robots. In: Siliciano, B. and Khatib, O. (eds.) Handbook of Robotics 2008. pp. 361–387. Springer (2008)
4. Fu, K. S., González, R. C. and Lee, C. S. G.: Robótica: control, detección, visión e inteligencia. McGraw-Hill, USA (1987)
5. Kelly, R. and Santibáñez, V.: Control de movimiento de robots manipuladores. Prentice Hall, Madrid, España (2003)
6. Nordin, M. and Frankel, V. H.: Biomecánica básica del sistema musculoesquelético. McGraw-Hill, España (2004)
7. Spong, M. W., Hutchinson, S., and Vidyasagar, M. : Robot Modeling and Control. John Wiley and Sons Inc., New York (2007)
8. Van der Loos, M. and Reinkensmeyer, D. J.: Reahilitation and health care robotics. In: Siliciano, B. and Khatib, O. (eds.) Handbook of Robotics 2008. pp. 1223–1246. Springer (2008)
9. Valdivia, C. H. G., Ortega, A. B., Salazar, M. A. O., Rivera, H. R. A.: Modelado y simulación de un robot terapéutico para la rehabilitación de miembros inferiores. Revista Ingeniería Biomédica, Vol. 7, No. 14, pp. 42–50 (2013).
10. The MathWorks: SimMechanics User's Guide, `http://www.mathworks.com`
11. Ollero Baturone, A.: Manipuladores robóticos y robots móviles. Alfaomega, España (2007)
12. Ma, Y., Hei, W. and Sam Ge, S.: Modeling and control of a Lower-Limb rehabilitation robot. ICSR 2012, Springer, pp. 581-590 (2012).

Design of a Biomechanical Model and a Set of Neural Networks for Monitoring of Weightlifting

Francisco Javier Rosas Ibarra, María Trinidad Serna Encinas,
César Enrique Rose Gómez, and Oscar Adrián Pinto García

Instituto Tecnológico de Hermosillo, División de Estudios de Posgrado e Investigación,
Maestría en Sistemas Industriales,
Av. Tecnológico S/N, Hermosillo, Sonora, México.

{frosas, tserna, crose}@ith.mx, elpintojr@gmail.com

Abstract. Nowadays, young athletes have injuries that leave them out the competitions, this injuries may occur due to the lack of the basic principles of weight training, muscle imbalance, fatigue, overhead, among others. The proposed weightlifting monitoring system is intended to coaches in order to provide them better control of the athletes during the training session, through continuous monitoring of exercise's measurements. This allows to detect certain risks of injury that may result if an exercise is poorly executed. To handle the uncertainty generated by the comparison of data obtained from athlete with the data of the biomechanical model for weightlifting, a set of neural networks to classify the athletes' movements during training is used. This paper describes the biomechanical model and the proposed neural network design for such monitoring system.

Keywords: Monitoring system, Weightlifting, Neural networks, Biomechanical model.

1 Introduction

In sports training, who is responsible for athletes to perform the right moves is the coach, but if he is in charge of several athletes at the same time, he may not notice that someone is performing wrong an exercise, or may fail to distinguish when an exercise is being poorly executed.

The monitoring system, that is sought to implement, will be used with elite/professionals athletes of the Sonora Sports Commission (CODESON), the system take measurements of the exercises carried out in weightlifting, using Kinect® sensors in a controlled environment. These measurements will be compared with a biomechanical model for weightlifting defined with specific movements for this exercise, and will be classified by a set of neural networks to determine if the positions of the executed exercise correspond to the positions of the base pattern previously defined and which is the precision level of the movements.

This paper focuses on the design of the set of neural networks to the monitoring system for the weightlifting training. Section 2 briefly presents the biomechanical model and the set of parameters defined in each phase of the exercise; Section 3 describes the proposed design of the neural network; specifically, it shows in detail the neural network for Phase 1. Section 4 discusses about the expected results; and finally, Section 5 presents the conclusion of this research.

2 Biomechanics Model

Sports biomechanics arises as a result of the development of the biomechanics of physical exercise, and was created by P. F. Lesgaft in the second half of the nineteenth century, it is responsible for evaluating a sport activity in order to design training techniques, improve performance and avoid injury [1].
A biomechanical model is a structure that represents the relationship between the objectives of the skills and the factors that produce those skills [2].

In Figure 1 the biomechanical model for weightlifting is shown, which is divided into six phases: Phase 1 (Initial position), which takes into consideration the angles of the ankles, knees, trunk and neck; Phase 2 (First pull), in addition to taking the angles also analyzes the horizontal scroll bar, vertical speed, maximum height of the bar and time; in Phase 3 (Adjustment) only the vertical velocity and time is analyzed; Phase 4 (Second pull) analyzed the same variables that Phase 2; Stage 5 (Slip) takes into account the time and bar height that it can reach in different types of lifts; for instance, slip unsupported (snatch) or slip supported (clean and jerk); Finally, in Phase 6 (Recovery), recovery height and position of the torso, arms and legs is analyzed [2-3].

2.1 Biomechanical Model with Parameters

In Figure 2 the biomechanical model for weightlifting with the parameters defined in each phase is observed. For Phase 1 the ankle angles must be 50°, the knee angles should be 72°, the trunk angle should be 45° and the neck angle should be 160°.

In Phase 2 ankle angles should be 25°, the knee angles should be 149°, the trunk angle should be 79° and the neck angle should be 160°; the horizontal scroll bar should be between -0.03 and -0.13 m; vertical velocity should be 1 to 1.6 m/s; the maximum height of the bar should be 28% to 32% of the height of the athlete; and time should be between 0.35 and 0.64 seconds.

For Phase 3, the vertical velocity should be between 0.88 and 1.52 m/s; and time should be between 0.08 and 0.19 seconds [2-5].

In phase 4 the angles of the ankle must be 32°, the knee angles should be 163°, the trunk angle must be 170° and 175° the angle of the neck; the horizontal scroll bar must be between 0 and -0.11 m; the vertical velocity should be 1.6 to 2.5 meter/second; the maximum height of the bar should be 60% to 66% of the height of the athlete, and time should be between 0.11 and 0.21 seconds.

For Phase 5, the time should be between 0.25 and 0.63 seconds; the height of the bar in the slip without support should be between 62% and 75% of the athlete's

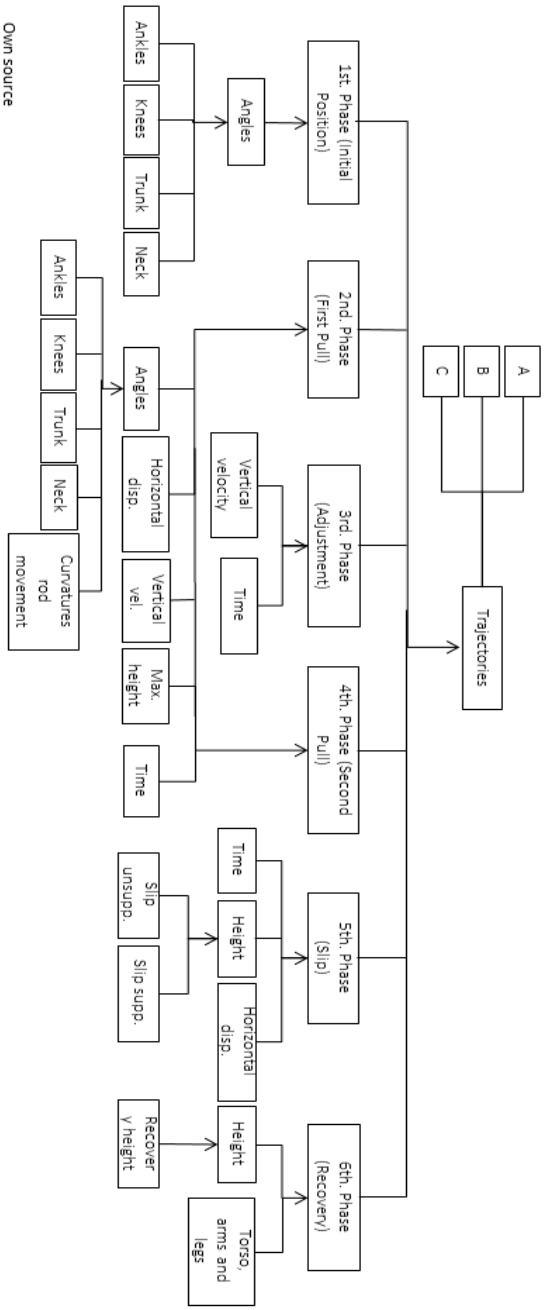


Fig. 1. Biomechanical model for weightlifting composed of 6 phases.

height, and bar height to support the slip should be between 75% and 95% of the athlete's height, and the horizontal scroll bar should be between -0.02 and -0.10 m.

For Phase 6 height recovery should be 106% and the position of the torso, arms and legs should be 180°.

2.2 Definition of Ranges

The definition of ranges is carried out taking into account the coaches, and even though they are aware that there is no perfect movement, it is desirable that athletes come close to it frequently. This was the reason that forced define ranges for each of the parameters that are in the biomechanical model previously described, and that are used by the set of neural networks to classify movements performed during training. The ranges used are those that occur especially in the angles and the degrees of freedom to allow in motion, so that it remains acceptable performance.

For example, in the knee angles of Phase 1, the parameter specifies 72°. The stated range for this parameter is ± 3°, i.e., the optimal motion varies between 69° and 75°, the acceptable movement is between 66° and 69° or between 75° and 78°, and the movement that is not acceptable is below 66° and above 78°.

3 Artificial Neural Network

An artificial neural network (ANN) is a set of simple units called neurons connected together by means of connections, and each element works only with local information and asynchronously. The knowledge is stored in the connections or synaptic weights, these weights will be adjusted according to the patterns that are available through a learning rule, this part of the process, in which the ANN learns the patterns, it is known as training. The knowledge acquired in this phase is reflected in certain values of the synaptic weights, with which the network must be able to generalize; that is, to give the correct output to a given input previously unseen [6-9].

Figure 3 shows the hierarchical structure of an ANN based system which begins with the neuron as is shown, then the layer is, which is a set of neurons in the same level, then the network itself is, a set of layers, and finally, the neuronal system containing inputs, a network, an algorithmic part and an output layer.

3.1 Design of the Neural Network Proposed to Monitoring System

The set of neural networks are designed based on the phases of the biomechanical model for weightlifting, that is, six neural networks, one for each phase, to allow classification of the movement, considering the ranges defined for each parameter as outputs of the set of networks.

3.2 Creation of the Neural Network

Figure 4 shows the neural network of Phase 1 that includes six neurons in the input layer, six neurons in the hidden layer, and six neurons in the output layer. For the

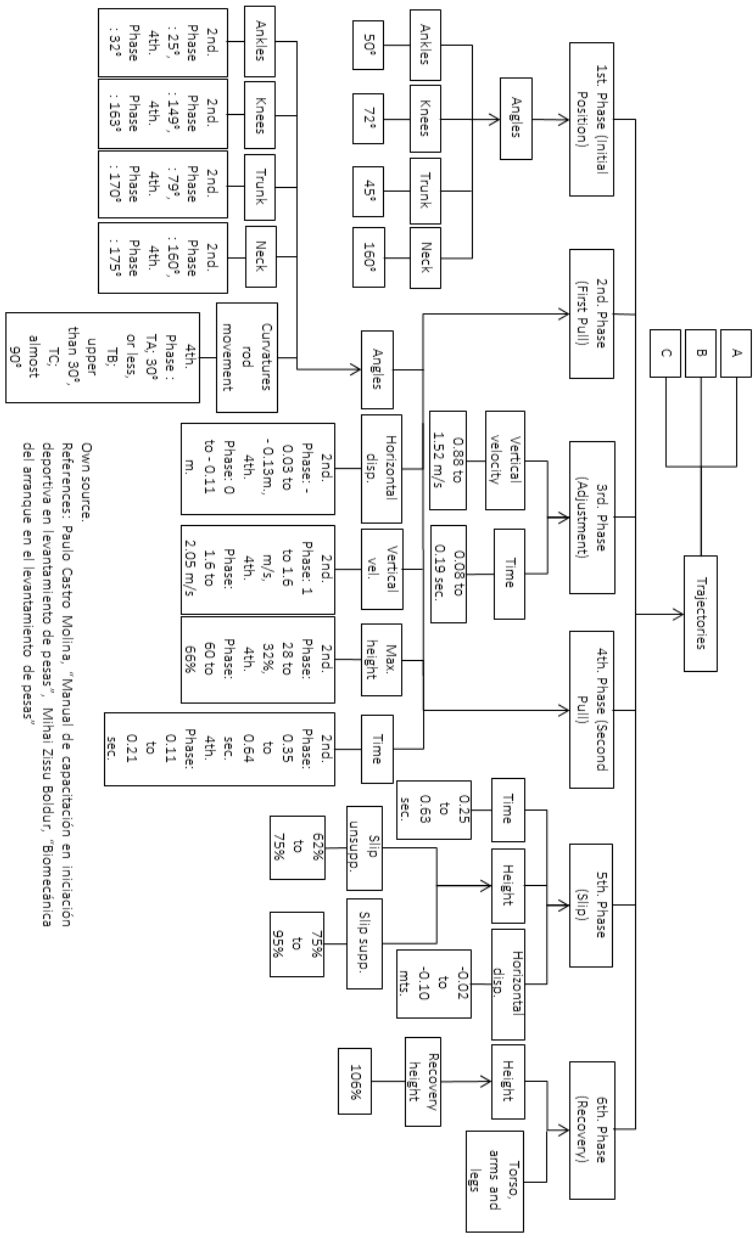


Fig. 2. Biomechanical model for weightlifting with parameters.

inputs we consider the angles of right and left knee (2 inputs), the right and left ankle (2 inputs), trunk (1 input) and neck (1 input), which make up the six input neurons.

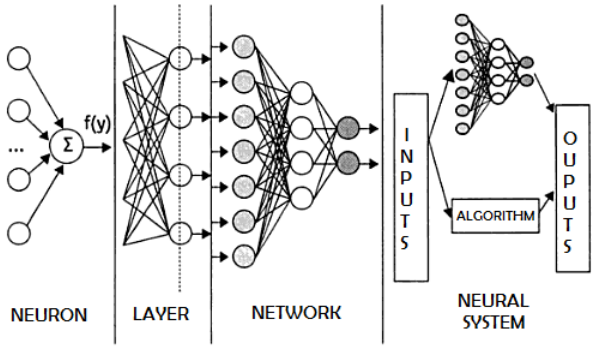


Fig. 3. Hierarchical structure of an ANN based system.

For each input has an output; i.e., an output for the right knee, one for the left knee, one for the right ankle, one for the left ankle, one for the trunk, and one for the neck, giving the six expected outputs.

Figure 5 shows the neural networks of Phase 2 and Phase 4 that includes eleven neurons in the input layer, eleven neurons in the hidden layer, and eleven neurons in the output layer. Besides consider the six angles mentioned before as inputs, it also has the bend of the bar (1 input), the horizontal scroll (1 input), the vertical speed (1 input), the maximum bar height (1 input) and time (1 input), which make up the eleven input neurons. Also it has an output for each input in Phase 2, making eleven expected outputs.

The neural network for Phase 3 only considers two inputs, which are vertical velocity (1 input) and time (1 input). It has two outputs, one for each input, giving the two expected outputs, has shown in figure 6.

Figure 7 shows he neural network for Phase 5 that considers three inputs, time (1 input), maximum bar height (1 input) and horizontal scroll (1 input). For each input there is an output, so there are three outputs on Phase 5.

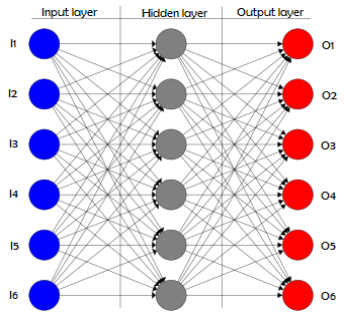


Fig. 4. Design of the neural network for the phase 1 (Initial position).

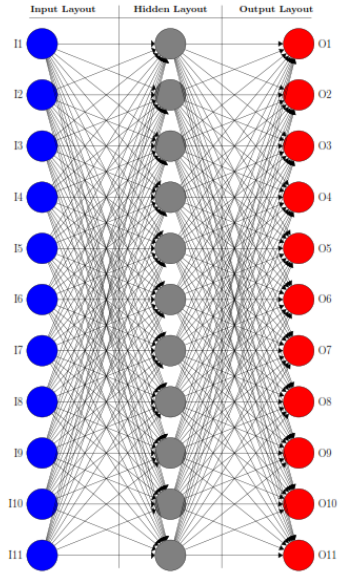


Fig. 5. Design of the neural network for phases 2 and 4 (First and second pull).

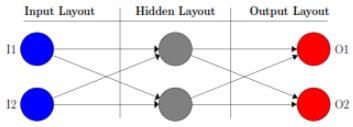


Fig. 6. Design of the neural network for the phase 3 (Adjustment).

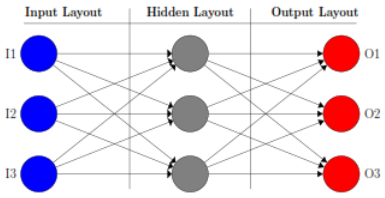


Fig. 7. Design of the neural network for the phase 5 (Slip).

Figure 8 shows the neural network of Phase 6 that has six inputs, the left and right knee angle (2 inputs), the trunk angle (1 input), the neck angle (1 input), the horizontal scroll (1 input) and the maximum bar height (1 input). There are six outputs in this neural network, one for each input.

The inputs of each phase correspond to the parameters set in each of the phases of the previously described model.

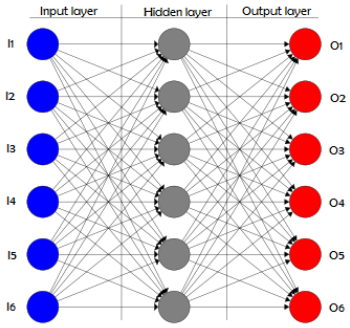


Fig. 8. Design of the neural network for the phase 6 (Recovery).

The activation function used in the neural network is set bipolar sigmoidal with alpha value equal to 0.5. Next is shown the part of the code that creates the neural network by assigning the number of inputs, outputs, the activation function and for neurons in the hidden layer.

```
this.m_networkSchema = new NetworkSchema(inputColumns.ToArray(),
outputColumns.ToArray(), cols);

IActivationFunction activationFunction = new
BipolarSigmoidFunction((double)sigmoidAlphaValue);

m_networkContainer = new NetworkContainer("Phase_1",
m_networkSchema, activationFunction, 6);
```

3.3 Training the Neural Network

The really important feature of neural networks is the ability to learn and adapt to the environment. The training process that makes that a neuron learn is vital part of the system and determines the degree of intelligence that can acquire the neural network [7].

For training the neural network of Phase 1 was required to record the movements of several athletes, from a very refined technique athlete, which performed the best moves according to the coach, and with another athlete with a less refined technique and with some mistakes.

In Table 1 are shown some of the input and output data used for training the neural network of Phase 1, which can be seen as inputs that have LeftAnkle_Angle, RightAnkle_Angle, LeftKnee_Angle, RightKnee_Angle, Neck_Angle and CenterHip_Angle; and outputs are O1, O2, O3, O4, O5 and O6. The outputs are taken at a range of 1 to 5, for example, for knees the 1 is angle less than 66° (not acceptable), a 2 is between 66° and 69° (acceptable), a 3 when it is between 69° and 75° (optimal), a 4 is between 75° and 78° (acceptable), and finally, 5 is an angle greater than 78° (not acceptable) angle.

In Table 2 are shown some of the input and output data used for training the neural network of Phase 1, which can be seen as inputs that have LeftAnkle_Angle, RightAnkle_Angle, LeftKnee_Angle, RightKnee_Angle, Neck_Angle, CenterHip_Angle, Curvature_Movement, Hor_Scroll, Vertical_Vel, Time and Maximum_Height.

Table 1. Input data and output data for the training of the neural network of the Phase 1 (Initial Position).

LAnkle Angle	RAnkle Angle	LKnee Angle	RKnee Angle	Neck Angle	CHip Angle	O1	O2	O3	O4	O5	O6
13.429	4.301	85.808	155.39	160.26	139.77	4	4	4	4	4	4
6.081	2.205	93.699	56.080	172.50	132.19	4	4	4	4	4	4
5.594	2.041	94.077	55.371	171.84	132.19	4	4	4	4	4	4
6.491	2.284	93.273	57.063	172.87	132.29	4	4	4	4	4	4
14.503	26.545	60.689	41.487	159.84	59.599	3	3	3	3	3	3
39.816	14.878	61.813	38.131	158.16	57.71	3	3	3	3	3	3
17.213	21.445	57.364	40.883	157.87	60.976	3	3	3	3	3	3
24.058	45.633	65.514	46.659	158.77	72.077	3	3	3	3	3	3
51.751	22.979	60.922	54.889	160.32	81.226	3	3	3	3	3	3

Table 2. Input data for the training of the neural network of the Phase 2 (First Pull).

LAnk Angle	RAnk Angle	LKnee Angle	RKnee Angle	Neck Angle	CHip Angle	Hor Scroll	Vert Vel	Time	Max Heig	Curv Mov
15.932	17.883	92.617	95.074	139.24	152.34	0.35	0.921	0.664	0.351	32.145
9.711	12.378	89.271	91.871	140.36	153.94	0.38	0.873	0.675	0.348	31.587
13.492	15.301	85.808	88.332	145.78	148.21	0.54	0.869	0.677	0.337	31.986
13.241	15.856	86.237	88.024	143.25	142.37	0.73	0.935	0.688	0.322	33.754
26.856	28.545	58.189	60.297	160.84	90.599	-0.01	1.12	0.38	0.283	23.541
31.189	34.578	59.093	61.031	159.16	87.071	0.01	1.28	0.39	0.287	28.152
41.997	43.743	63.574	65.753	158.97	87.976	0.03	1.35	0.45	0.292	27.378
48.723	51.683	65.348	66.659	158.13	78.573	0.08	1.42	0.47	0.301	29.961

Table 3. Output data for the training of the neural network of the Phase 2 (First Pull).

O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11
4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4
3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	3	3	3	3

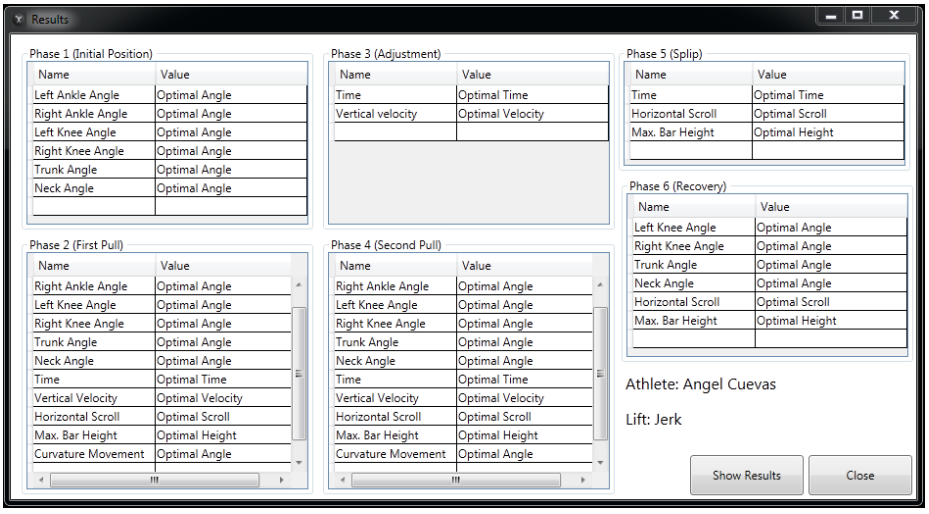


Fig. 9. Results obtained in jerk for an athlete with refined technique.

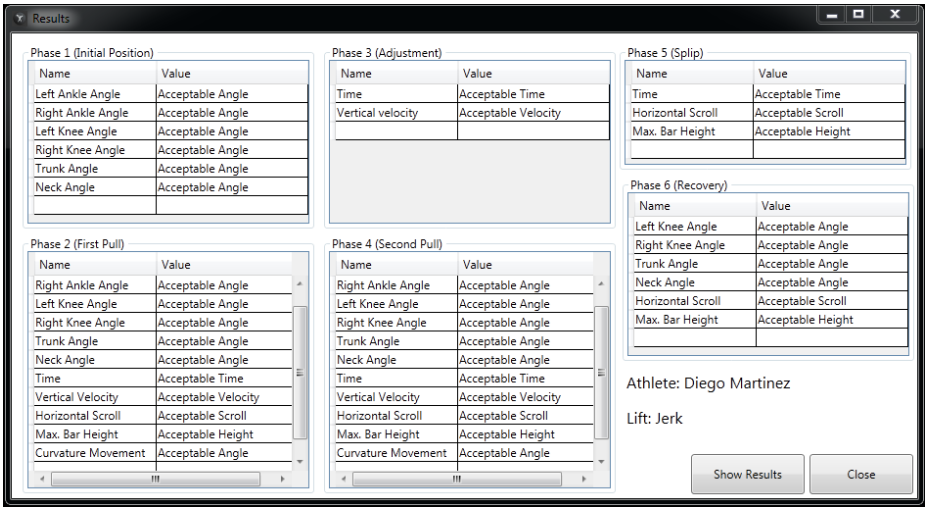


Fig. 10. Results obtained in jerk for an athlete with no refined technique.

In Table 3 are shown some of the outputs using a range of 1 to 5, for example, for horizontal scroll the 1 is scroll less than -0.18 (not acceptable), a 2 is between -.18 and -0.13 (acceptable), a 3 when it is between -.013 and 0.3 (optimal), a 4 is between 0.3 and 0.8 (acceptable), and finally, 5 is a scroll greater than 0.8 (not acceptable) scroll.

Phase 4 has the same input parameters than Phase 2, but with the values for the training of the Phase 4. The Phase 3 has 2 inputs parameters, the Phase 5 has 3 input

parameters and the Phase 6 has 6 input parameters, with the corresponding values of each phase.

For training the network a learning rate of 0.1, a momentum of 0.0, and the limit of error in 0.1 are used, the learning function is Backpropagation. Below is shown the code with which the network is trained:

```
BackPropagationLearning networkTeacher = new
BackPropagationLearning(m_networkContainer.ActivationNetwork);
networkTeacher.LearningRate = options.firstLearningRate;
networkTeacher.Momentum = options.momentum;

this.m_networkState.ErrorTraining =
networkTeacher.RunEpoch(options.TrainingVectors.Input,
options.TrainingVectors.Output)

if (m_networkState.ErrorTraining <= options.limError)
    stop = true;
```

4 Expected Results

By integrating the outputs of the set of neural network we can know the classification of movements performed by the athlete, also we can know the movements that must be corrected and the phases in which they were executed, therefore, we expected that this help to improve his technique and avoid risk of injury.

Figure 9 shows the results obtained after the use of the set of neural networks for jerk, you can see the 6 phases with all the parameters to be analyzed (right and left knees angles, right and left ankles angles, neck and trunk angles, time, vertical velocity, horizontal scroll, maximum bar height, curvature movement, depending on each phase) and outputs for each them.

Figure 10 shows, similarly, the results after the use of the set of neural networks for jerk, where the result of the parameters of the six phases to be analyzed for an athlete with no refined technique is deployed.

5 Conclusions

The biomechanical model proposed for weightlifting consists of six phases (initial position, first pull, adjustment, second pull, slip and recovery). Each phase consists of a set of parameters, which specify the optimal value to achieve during the performance of movements. With the literature, it is possible to define the optimum value of approximately half of the parameters; other values were determined with the support of the coaches. Similarly, it was possible to define the range of values for each of the (optimal, acceptable and not acceptable) parameters.

Using the proposed biomechanical model for weightlifting, where the set of parameters for each phase and the ranges for each parameter are defined, we obtain a design for the set of neural networks for the classification of movements performed

during the exercise, which will end system implementation and present the analysis of results.

Acknowledgment. We thank "Dirección General de Educación Superior Tecnológica" (DGEST) for the scholarship with number 704464, issued from August 2012 to July 2014 for the completion of a Masters in Industrial Systems.

References

1. Ford J., Ángel E.: La biomecánica deportiva en el arte marcial del full contact, Tesis de licenciatura, Universidad Central de Venezuela, Venezuela, (2010)
2. Beltrán, S. P., Boldur, M. Z.: Características cinemáticas de la fase del halón en el arranque, Levantamiento de pesas femenino. EAE Editorial Academia Española, ISBN-13: 9783846567777, p. 60, España, (2011)
3. Ford, A. E.: La biomecánica deportiva en el arte marcial del full contact: Un reportaje educativo e instruccional, Ph.D. dissertation, (2013)
4. Sobarzo, P. D.: Biomecánica deportiva, Boletín informativo de la Coordinación de la Investigación Científica, el faro, no. 88 - 89, p. 8 – 10, (Agosto 2008)
5. Suárez, G. R.: Biomecánica deportiva y control del entrenamiento, Primera edición, Funámbulos Editores, Medellín, Colombia, (2009)
6. Nigrin, A., Neural networks for pattern recognition. MIT Press, (1993)
7. Isasi P., Galván I.: Redes de neuronas artificiales un enfoque práctico, Pearson Education, ISBN: 84-205-4025-0, Madrid, (2004)
8. Del Brío, B. M., Molina, A. S.: Redes neuronales y sistemas difusos, (2002)
9. Matich, D. J.: Redes Neuronales: Conceptos básicos y aplicaciones, Cátedra de Informática Aplicada a la Ingeniería de Procesos–Orientación I, (2001)

Técnicas de reducción, algoritmos resistentes al ruido o ambos. Opciones para el manejo de rasgos clasificatorios en la atribución de autoría.

Antonio Rico Sulayes

Universidad Autónoma de Baja California, Calzada Universidad 14418, Parque Industrial Internacional, Tijuana, Baja California, México
antonio.rico@uabc.edu.mx

Resumen. Un problema en las tareas clasificatorias es el manejo de los rasgos que caracterizan las clases. Cuando la lista de rasgos es larga, se puede usar un algoritmo resistente al ruido de rasgos irrelevantes, o reducir dichos rasgos. La atribución de autoría, tarea que asigna un texto anónimo a un sujeto en una lista de posibles autores, ha sido ampliamente abordada como una tarea de clasificación automática de texto. En dicha tarea, los *n-gramas* pueden producir listas largas de rasgos incluso en corpus pequeños. A pesar de esto, falta una investigación que exponga los efectos de usar algoritmos resistentes al ruido, reducir los rasgos, o combinar ambas opciones. Este trabajo responde a esta carencia utilizando contribuciones a foros de discusión relacionados con el crimen organizado. Los resultados obtenidos muestran que los clasificadores evaluados, en general, se benefician con la reducción de rasgos, y que gracias a dicha reducción, incluso algoritmos clásicos superan a clasificadores de punta considerados altamente resistentes al ruido.

Palabras clave: Atribución de autoría, Rasgos clasificatorios, Algoritmos resistentes al ruido, Reducción de rasgos.

1 Listas de rasgos en la atribución de autoría

Si se define a la atribución de autoría como la asignación de un texto anónimo a un sujeto dentro de una lista de posibles autores, esta tarea constituye un problema de clasificación de textos. Ahora bien, si este problema se aborda utilizando métodos automatizados, el problema le compete a la clasificación automática de textos, área atendida por la recuperación de información. Como un problema de la clasificación automática de textos, la atribución de autoría utiliza dos elementos primordiales. Por un lado, requiere una selección de rasgos clasificatorios, que discriminan a los elementos de las diversas clases. Por otro lado, la atribución de autoría utiliza un método de clasificación que procesa los rasgos. En este contexto, el método clasificatorio se emplea para atribuir un cierto texto a un sujeto en específico.

Respecto del primer elemento esencial para la atribución de autoría, la selección de rasgos clasificatorios, los investigadores dedicados a dicha tarea han propuesto nue-

vos rasgos de manera constante por varias décadas. A finales del siglo pasado, más de 1,000 rasgos diferentes fueron identificados en más de 300 trabajos dedicados a esta tarea clasificatoria [1]. Este número de rasgos se ha incrementado dramáticamente en años recientes debido a la introducción de rasgos textuales que son etiquetados de manera automática, como por ejemplo, los n-gramas. Los n-gramas fácilmente producen listas de varios millares de elementos incluso en corpus (colecciones de textos de lenguaje natural) relativamente pequeños, como los que típicamente se utilizan en la atribución de autoría [2,3,4].

En contraposición a la proliferación de rasgos en la atribución de autoría, los resultados de investigaciones recientes surgieren que la selección de los mismos es el elemento primordial en la mejora de resultados para esta tarea [5,6,7]. Según las investigaciones mencionadas, esta selección es incluso más importante que la puesta a punto de los algoritmos de clasificación. En la clasificación de textos en general, la motivación para el uso de listas reducidas de rasgos obedece a que los rasgos altamente discriminatorios son más eficientes y obtienen una mayor precisión en los resultados [8]. Estos rasgos discriminatorios evitan el ruido de las listas extensas, las cuales incluyen rasgos redundantes o poco discriminatorios. Este tipo de rasgos es particularmente ineficiente cuando se aplica a nuevos conjuntos de datos. Una respuesta que se ha dado a este problema en la atribución de autoría son los algoritmos avanzados para la clasificación de textos, tales como las máquinas de vectores de soporte (SVMs, por sus siglas en inglés), que pueden compensar el ruido de las largas listas de rasgos [3], [9]. La segunda respuesta es el uso de técnicas para la reducción de dichas listas, como la selección de rasgos con mayor frecuencia o con índices altos de información mutua. Esta solución también ha sido extensamente empleada en la atribución de autoría [10]. Frente a estas dos posibilidades en el manejo de rasgos clasificatorios para la atribución de autoría (la utilización de algoritmos que compensan el ruido de los rasgos no discriminatorios o redundantes y la reducción alternativa de las listas de rasgos con recursos externos al clasificador), la literatura especializada no ha comparado los resultados de ambas opciones.

Este artículo responde a esta carencia, comparando los métodos de clasificación y de reducción de rasgos más comunes en la atribución de autoría. Además este trabajo introduce un método de reducción de rasgos nunca antes usado en esta tarea. En la evaluación de las diferentes combinaciones entre clasificadores y técnicas de reducción de rasgos, el presente trabajo utiliza datos extraídos de medios sociales relacionados con el crimen organizado en México. El artículo concluye mostrando que los métodos de clasificación con una larga tradición en la atribución de autoría [1] se pueden combinar con técnicas de reducción de rasgos (tanto técnicas conocidas como de nueva aplicación en este contexto) y que esas combinaciones igualan y superan los resultados obtenidos por clasificadores de última generación.

2 Los medios sociales y el crimen organizado

La guerra contra el tráfico de drogas en México se caracterizó desde un inicio por un creciente y luego constante número de muertes [11,12,13,14]. Además de esta característica, esta lucha también se ha distinguido por el continuo uso de comunicaciones escritas enviadas por miembros del crimen organizado tanto al gobierno como

a la población en general [7]. Entre estas comunicaciones se encuentran los comentarios publicados por usuarios en medios sociales dedicados a este tema. Este tipo de medios sociales han proliferado en años recientes en este país, con algunos resultados lamentables, como el asesinato de sus usuarios por grupos criminales y el eventual cierre de algunos de estos sitios [15,16,17]. A pesar de estos hechos, este tipo de medios sigue existiendo hasta el día de hoy [18].

El presente trabajo utiliza las contribuciones de usuarios publicadas en uno de los primeros sitios relacionados con el crimen organizado en México. Este sitio, creado en abril de 2010, albergó en su inicio un foro de discusión dedicado a este tema [7]. Las contribuciones de este foro utilizadas aquí fueron recuperadas copiando todos los mensajes publicados durante el primer medio año de vida del foro. Esto permitió recuperar 41,751 mensajes publicados en 4,205 conversaciones. Una vez depurados todos los mensajes recuperados (eliminando copias y mensajes de usuarios anónimos) se identificaron 37,571 mensajes creados por usuarios registrados en el foro. Estos mensajes pertenecen a 1,026 usuarios diferentes y contienen un total de 2,128,049 instancias de palabras o tokens.

2.1 Datos experimentales

Con los datos recuperados del foro de discusión referido arriba, se crearon varios corpus para explorar los efectos de la reducción de rasgos en combinación con diversos clasificadores, comunes en la atribución de autoría. Entre los 1,026 usuarios que produjeron mensajes usando una cuenta de usuario, se seleccionaron aquellos que tenían un mínimo de 40 mensajes individuales con un mínimo de 2,000 palabras de texto original en la suma de todos sus mensajes. Con estos dos criterios de selección, se identificaron 106 usuarios del foro que cumplían con dichos criterios.

El mínimo de palabras de texto original que se puso como requisito para seleccionar a los usuarios del foro (2,000 palabras) se encuentra en el extremo inferior de lo que estudios previos han utilizado en la atribución de autoría. Por ejemplo, entre los investigadores que reportan este dato experimental algunos han empleado 2,000, 8,000, 15,000, 33,000, 40,000, y 55,000 palabras, [19], [20], [21], [22], [23], [2], respectivamente. Esta cantidad de texto es utilizada como datos de entrenamiento para representar a cada sujeto en el conjunto de autores potenciales durante la clasificación. El otro criterio de selección de sujetos como autores potenciales (un mínimo de 40 mensajes), se utilizó para descartar usuarios esporádicos con escasos mensajes de cierta extensión. Muestreando aleatoriamente los 106 seleccionados, se identificaron 40 usuarios con los que se construyeron 39 corpus. En estos corpus, el número de sujetos va de un mínimo de 2 hasta un máximo de 40.

Dividiendo las 2,000 palabras de texto original de cada sujeto, se construyeron 4 sub-muestras de aproximadamente 500 palabras cada una. Estas sub-muestras fueron construidas agregando aleatoriamente mensajes, de entre todas las contribuciones de cada uno de los usuarios del foro, conservando la integridad de los mensajes individuales. Cada una de estas sub-muestras se utilizó como una unidad en los datos de prueba empleados para la clasificación. El rango en el tamaño de estas sub-muestras de prueba (478-541 palabras) se encuentra también en el extremo inferior de lo que

estudios previos que utilizan mensajes íntegros han empleado. Por ejemplo, algunos estudios usan 99-608, 100-1000, 500-2000, 600, 628-1342, y 7,500+ palabras en las unidades de sus datos de prueba, [19], [24], [22], [23], [21] y [20], respectivamente. Debido a que las contribuciones de cada usuario seleccionado fueron agregadas en 4 sub-muestras, los experimentos realizados con cada uno de los 39 corpus comprendieron desde 8 atribuciones (en el corpus más pequeño con 2 autores) hasta 160 atribuciones (en el corpus más grande con 40 sujetos).

3 Atribución de autoría como clasificación automática de textos

La atribución de autoría ha sido ampliamente abordada, tanto por investigadores de la recuperación de la información, así como de la lingüística y la estilística forense [5]. En todas estas disciplinas, establecer rasgos de comparación y clasificación entre los textos de diversos autores es una necesidad esencial. Sólo a partir de estos rasgos se pueden discriminar los textos, los cuales son tratados como anónimos durante la puesta a prueba del método de clasificación. Ahora bien, en el caso particular de la recuperación de información, el procesamiento de los rasgos mencionados es realizado siempre de manera automática con un algoritmo clasificador. Además, algunos de los clasificadores comunes en esta área, los cuales son considerados de última generación, han surgido dentro del aprendizaje automático.

3.1 Rasgos clasificatorios de autoría

En cuanto a los rasgos que se utilizaron para realizar la clasificación, se partió de una selección previa de rasgos léxicos, sintácticos y estructurales. Los rasgos léxicos incluyeron una lista de todos los unigramas de palabras (equivalente a todos los tipos o formas léxicas diferenciadas). Como esta lista es dependiente del corpus del que es extraída, el tamaño de la misma varió entre 1,402 tipos para el corpus más pequeño, hasta 13,089 para el corpus más grande. Cabe aclarar que la puntuación fue removida de las unidades léxicas a las que se unía y los signos separados fueron utilizados como unigramas léxicos independientes, un procedimiento común en la atribución de autoría [9], [20], [25] y [26]. En cuanto a los rasgos sintácticos, se utilizó una lista previamente recabada (para otra tarea clasificatoria) con elementos léxicos funcionales pluriverbales, es decir, con más de una palabra. Estos elementos se componen principalmente de una preposición más otros elementos léxicos, como ‘después de(l)’ o ‘lejos de(l)’, o de una conjunción combinada con otras palabras, como ‘después de que’ o ‘mientras que’. La lista predeterminada de elementos léxicos funcionales pluriverbales, cuyas instancias fueron etiquetadas en los corpus, tenía 132 elementos en total, con 68 bigramas, 56 trigramas, and 7 tetragramas. Finalmente, los rasgos estructurales estaban dados por una lista preseleccionada de 19 elementos, varios de ellos previamente utilizados por el autor de este artículo [7]. Entre estos rasgos se incluyen diversos rasgos de formato del texto como el uso de subrayados, negritas, imágenes, colores y tamaños especiales de letras. Igualmente, estos rasgos comprenden elementos propios de la comunicaciones electrónicas, como el uso hipervínculos, activos y

no activos, emoticones, en imágenes y representados con caracteres del teclado, así como la reduplicación excesiva de signos de puntuación, como suelen emplearse los signos de admiración. Dada esta preselección de rasgos de autoría, el número de rasgos etiquetados (todos automáticamente) en los 39 corpus va de 1,553 para el corpus con 2 autores solamente, hasta 13,249, para el corpus con el máximo de 40 sujetos.

3.2 Algoritmos de clasificación en la atribución de autoría

Las investigaciones sobre atribución de autoría han empleado una variedad de algoritmos clasificatorios. En una revisión exhaustiva de 32 trabajos dedicados a esta tarea y publicados en la última década, se identificaron 23 algoritmos diferentes de clasificación [10]. Aunque muchos de los clasificadores identificados aparecen sólo una vez en la literatura, algunos algoritmos han sido utilizados en varias investigaciones. Entre estos algoritmos se encuentran diferentes implementaciones del árbol de decisión, C4.5, formas varias del análisis bayesiano (como el multivariante y el basado en el modelo de Bernoulli), diferentes tipos de redes neuronales (como las artificiales y las llamadas redes neuronales de retropropagación [9]) y las SVMs. También son comunes en la atribución de autoría algunos clasificadores propios de la estadística, como el análisis discriminante (AD) y los clasificadores basados en la prueba de Chi cuadrado. Además, hay que mencionar que 10 de los 32 trabajos revisados en el estudio citado utilizan más de un algoritmo clasificador y comparan los resultados obtenidos por los diferentes algoritmos seleccionados.

En cuanto a los clasificadores puestos a prueba en este trabajo, se escogieron los 4 algoritmos que han dado los mejores resultados en la atribución de autoría, según la revisión bibliográfica exhaustiva antes mencionada [10]. Los 4 clasificadores son el AD, el análisis bayesiano multivariante (ABM), el análisis bayesiano de Bernoulli (ABB) y las SVMs. A estos clasificadores se ha sumado el algoritmo más comúnmente usado como base de referencia en esta tarea, el árbol de decisión C4.5, en su implementación para Weka, J4.8.

3.3 Técnicas de reducción de rasgos en la atribución de autoría

El estudio mencionado en la sección anterior [10], reporta también un abundante uso de técnicas de reducción de rasgos en las investigaciones dedicadas a la atribución de autoría. De los 32 estudios examinados, 17 fueron identificados por su uso de alguna técnica de reducción de rasgos, o de algún método de evaluación de rasgos que permite reducir el conjunto de los mismos. Con menos variación que el uso de clasificadores, las técnicas de reducción de rasgos identificadas incluyen la ganancia de información (GI), la frecuencia (relativa, absoluta o normalizada), el análisis de componentes principales, algunos métodos de evaluación de rasgos de la estadística en general (el análisis de la varianza, ANOVA, el análisis de la covarianza, ANCOVA, y la ANOVA de dos vías), y dos métodos paso a paso, la distancia de Mahalanobis y la Lambda de Wilks. Un estudio también utiliza una lista de palabras vacías para eliminarlas de su lista completa de rasgos. Aunque sólo dos estudios comparan más de una

técnica de reducción, el de uso de dos de ellas es notoriamente más común que el de las demás: la frecuencia, empleada en seis estudios, y la GI, en tres.

En este trabajo, se escogieron las dos técnicas de reducción más comunes, la frecuencia y la GI, que son también aquellas con que se han reportado los mejores resultados en los estudios comparativos. La frecuencia aquí ha sido expresada como frecuencia absoluta con un mínimo de 4 instancias, número igual al número de sub-muestras por autor. Además se decidió incluir una técnica de reducción nueva en la atribución de autoría, la llamada "selección de sub-conjuntos de rasgos basada en correlaciones" (CFS, por sus siglas en inglés). Esta técnica de reducción, descrita por primera vez en [27], se incluyó porque fue diseñada con la intención explícita de mejorar el rendimiento de algoritmos basados en el análisis bayesiano. Como se mencionó arriba, dos versiones de estos algoritmos, ABM y ABB, fueron utilizados aquí.

Por otro lado, la lista completa de rasgos sin reducción constituyó la cuarta alternativa para combinar con cada algoritmo de clasificación. En esta cuarta opción es trabajo exclusivo del clasificador compensar el ruido de los rasgos no discriminatorios o redundantes. Dados los cinco clasificadores y las cuatro listas de rasgos (tres reducidas y una sin ninguna reducción), hay un total de 20 configuraciones aplicadas a los 39 corpus. Todas estas opciones dan origen a un total de 780 experimentos.

4 Resultados

Respecto de la precisión cuyos promedios se reportan a continuación, ésta representa la proporción de verdaderos positivos o asignaciones correctas de sub-muestras de prueba a sus verdaderos autores. Para la obtención de dicha precisión en experimentos individuales, se aplicaron un clasificador y una lista de rasgos (reducida o sin ninguna reducción) en la asignación de todas las sub-muestras de prueba a su respectivos autores. Además, el cálculo de la precisión en los experimentos individuales se llevó a cabo por medio de un diseño de validación cruzada. En cuanto a las figuras presentadas a continuación en la Tabla 1, éstas representan el promedio de la precisión obtenida por la combinación de un clasificador y una lista de rasgos al aplicarse paulatinamente a los 39 corpus.

Tabla 1. Resultados promedio de la clasificación en los 39 corpus

Técnica de Reducción	Clasificador				
	C4.5	AD	ABM	ABB	SVMs
Ninguna	0.456	0.313	0.743	0.831	0.457
Frecuencia	0.489	0.466	0.942	0.821	0.829
GI	0.671	0.700	0.947	0.848	0.775
CFS	0.660	0.811	0.940	0.820	0.726

En su conjunto, la Tabla 1 muestra los resultados en promedio, sobre todos los 39 corpus, de los 5 clasificadores en combinación con las 4 listas de rasgos. Como se puede observar en la tabla, los mejores resultados obtenidos con el algoritmo de base de referencia, C4.5, se logran al combinar el mismo con la técnica de reducción por GI. Lo mismo sucede con los clasificadores basados en análisis bayesianos, el ABM y el ABB. En los tres clasificadores que obtienen sus mejores resultados al combinarse con la lista reducida por GI, la precisión promedio con todos los corpus incrementa de 0.456 sin ninguna reducción a 0.671 para C4.5, de 0.743 a 0.947 para el ABM, y de 0.831 a 0.848 para el ABB. El mejor resultado en promedio sobre todos los corpus para el clasificador del AD se da con la técnica de reducción de CFS. Con esta técnica, la precisión promedio mejora de 0.313 sin ninguna reducción a 0.811. Finalmente, con el clasificador de última generación, basado en SVMs, el mejor resultado en promedio se obtiene utilizando la frecuencia absoluta mínima de 4 instancias. Con este criterio, la precisión promedio mejora de 0.457 sin ninguna reducción a 0.829.

5 Conclusiones

Con la perspectiva global que ofrece la Tabla 1, es posible regresar a la lista de opciones para el manejo de rasgos clasificatorios mencionada en el título de este trabajo. De esa lista se desprende la pregunta: ¿Cuál es la mejor opción para manejar listas largas de rasgos en la atribución de autoría: aplicar técnicas de reducción de rasgos, usar algoritmos de clasificación resistentes al ruido o combinar estas dos opciones? En cuanto a la primera respuesta a la pregunta, la aplicación de técnicas de reducción de rasgos, la Tabla 1 muestra que todos los clasificadores se benefician claramente de dicha aplicación. Al mismo tiempo, se aprecian tres tendencias en el beneficio obtenido por medio de las diversas técnicas de reducción. Por un lado, dos clasificadores, C4.5 y el AD, observan una mejoría moderada con la lista reducida por medio de la frecuencia. La mejoría va de 0.456 a 0.489 para C4.5 y de 0.313 a 0.466 para el AD. Sin embargo, la mejora en la precisión de estos dos clasificadores es mayor con las técnicas que requieren un cálculo más elaborado y que, probablemente de manera más relevante, producen listas de rasgos mucho más cortas. En este sentido cabe mencionar que el número de rasgos en la lista reducida con el criterio de frecuencia va de 229 en el corpus con dos autores a 2,168 en el corpus con 40 autores, mientras que GI y CFS generan listas considerablemente más pequeñas, con rangos de 23-75 rasgos y de 7-27 rasgos, respectivamente. La segunda tendencia que se observa en la aplicación de técnicas de reducción está dada por otros dos clasificadores, el ABM y las SVMs. Estos algoritmos se benefician notoriamente con todas las técnicas de reducción de rasgos. Finalmente, el ABB exhibe una tercera tendencia ya que no siempre se beneficia con la reducción de rasgos. Sin embargo, logra una mejoría moderada, respecto de la lista sin reducción, al alimentarse con la lista reducida por GI.

A continuación, la Figura 1 muestra la precisión obtenida por la combinación de la lista de rasgos seleccionados por GI y el ABM. Esta combinación obtuvo la precisión promedio más alta sobre todo los corpus. En la misma figura, se ha incluido la precisión obtenida por el ABM sin ninguna reducción de rasgos.

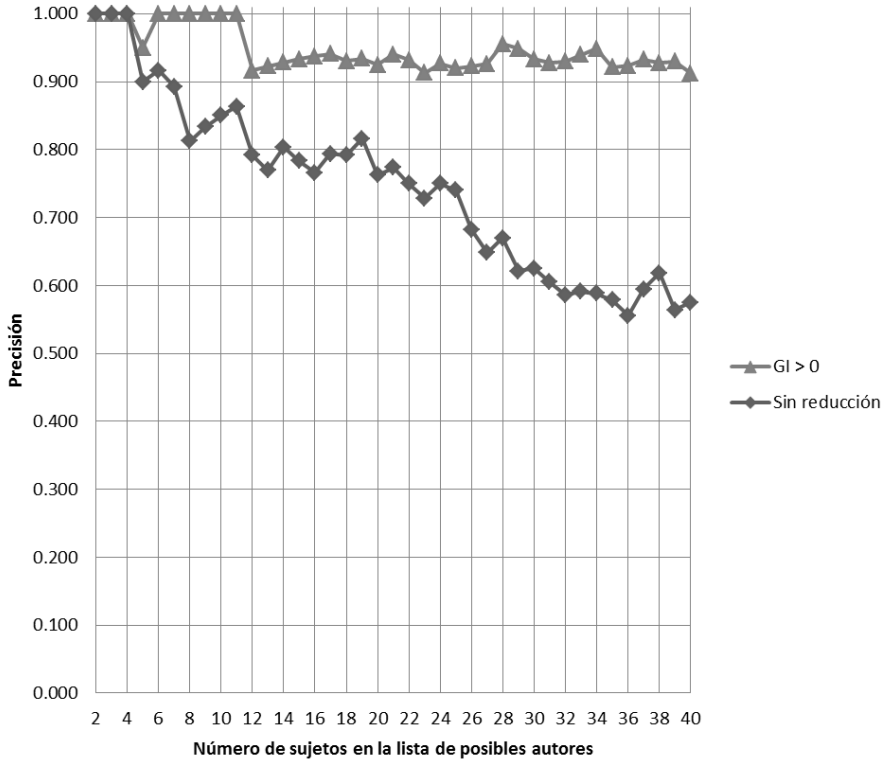


Fig. 1. Efecto de la reducción de rasgos por GI con el ABM

La línea superior en la Figura 1, con vértices resaltados con triángulos, corresponde a la precisión del ABM con la lista reducida con GI, mientras que la línea inferior, con vértices resaltados con rombos, describe la precisión del ABM al usar la lista no reducida. Estas dos líneas permiten apreciar el claro efecto positivo de la reducción de rasgos en la combinación más exitosa, en este estudio, de una técnica de reducción y un clasificador. Como se observa en la línea superior, la aplicación de la técnica de reducción de rasgos por GI le permite al clasificador del ABM mantener una precisión constante por encima de 0.900 a lo largo de los todos los experimentos con los 39 corpus, los cuales incluyen como muestra la figura desde dos hasta 40 autores.

Respecto de la segunda opción para manejar rasgos clasificatorios, la utilización de clasificadores resistentes al ruido, este trabajo ha evaluado un algoritmo de última generación en el aprendizaje automático, las SVMs. Este algoritmo es considerado como especialmente resistente al ruido en el contexto de la atribución autoría [3], [9]. En este estudio, las SVMs han mostrado por un lado que son en efecto altamente resistentes al ruido, ya que obtienen su mejor resultado promedio con las listas reducidas con el criterio de frecuencia, las cuales son comparativamente largas. Sin embargo, la combinación de un modelo tradicional de clasificador, el ABM, en combinación

con cualquiera de los tres métodos de reducción de rasgos (frecuencia, GI y CFS), tiene la capacidad de superar los mejores resultados obtenidos por las SVMs. La superioridad del ABM, en su mejor desempeño logrado con GI, por encima del mejor rendimiento de las SVMs, al combinarse con la lista reducida por frecuencia, puede apreciarse en la Figura 2, a continuación.

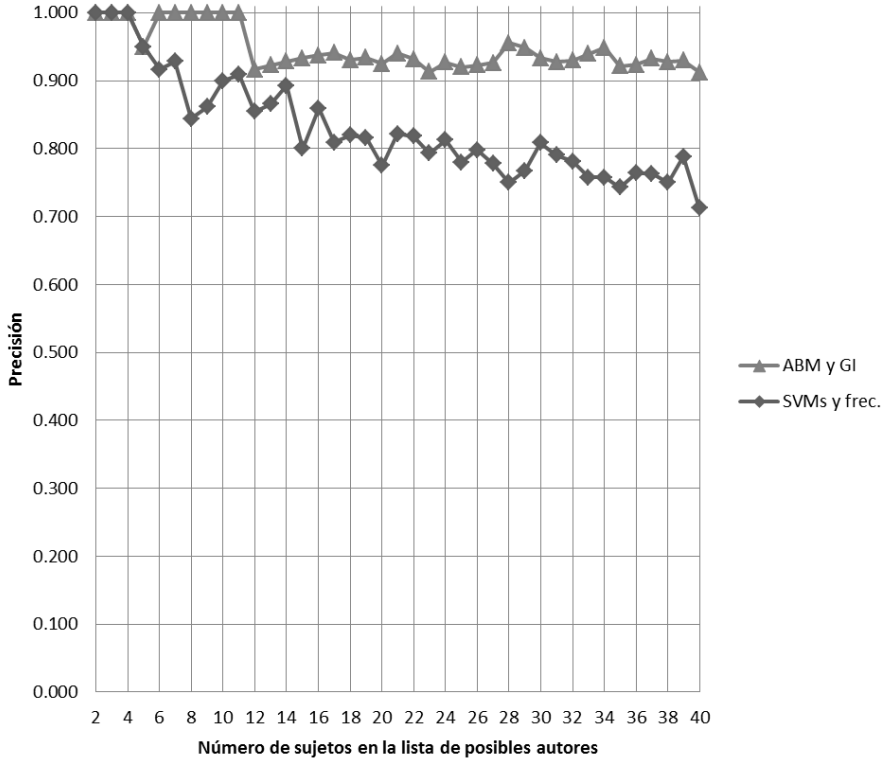


Fig. 2. Precisión promedio más alta obtenida por el ABM y las SVMs

La tercera opción para el manejo de listas largas de rasgos es la utilización simultánea de algoritmos altamente resistentes al ruido y técnicas de reducción. La conveniencia de utilizar o no esta opción se desprende de los últimos dos puntos discutidos. Por un lado, todos los clasificadores se beneficiaron de las técnicas de reducción de rasgos. Por otro lado, el algoritmo altamente resistente al ruido demostró que obtuvo el mayor beneficio al usar listas reducidas comparativamente largas. Sin embargo, ha sido la combinación de un clasificador que es no es considerado particularmente resistente al ruido y una lista notoriamente reducida (y depurada de aquellos rasgos que insertan ruido) la que ha obtenido los mejores resultados en la evaluación de las múl-

tiples combinaciones de clasificadores y técnicas de reducción. Por tanto, esta tercera opción ha sido superada, al igual que la segunda antes discutida.

Referencias

1. Rudman, J. The State of Authorship Attribution Studies: Some Problems and Solutions. Computers and the Humanities. 31, 351-365 (1998)
2. Koppel, M., Schler, J., & Argamon, S. Computational Methods in Authorship Attribution. Journal of the American Society for Information Science and Technology. 60(1), 9-26 (2009)
3. Abbasi, A., & Chen, H. Applying Authorship Analysis to Extremist-Group Web Forum Messages. IEEE Intelligent Systems. 20(5), 67-75 (2005)
4. Gamon, M. Linguistic Correlates of Style: Authorship Classification with Deep Linguistic Analysis Features. En actas de the 20th International Conference on Computational Linguistics: Vol.4, pp. 611-617. Stroudsburg, PA: Association for Computational Linguistics (2004)
5. Juola, P. Authoship Attribution. Hanover, MA: Now Publishers (2008)
6. Koppel, M., Schler, J., & Messeri, E. Authorship Attribution in Law Enforcement Scenarios. En C.S. Gal, P. Kantor, & B. Saphira (Eds.), Security Informatics and Terrorism: Patrolling the Web, pp.111-119. Amsterdam: IOS (2008)
7. Rico-Sulayes, A. Statistical Authorship Attribution of Mexican Drug Trafficking Online Forum Posts. International Journal of Speech, Language and the Law. 18(1), 53-74 (2011)
8. Manning, C. D., Raghavan, P., & Schütze, H. Introduction to Information Retrieval. New York, NY: Cambridge (2008)
9. Zheng, R., Li, J., Chen, H. & Huang, Z. A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques. Journal of the American Society for Information Science and Technology. 57(3): 378-393 (2006)
10. Rico-Sulayes, A. (2012). Quantitative Authorship Attribution of Users of Mexican Drug Dealing Related Online Forums. Tesis doctoral, Georgetown University.
11. Agar, M. Mexican Drug War Deaths Top 47,500. The Telegraph (2012, enero 12)
12. Córdoba, J., & Luhnow, D. In Mexico, Death Toll in Drug War Hits Record. The World Street Journal (2011, enero 13)
13. Davison, J., & Stastna, K. Mexico's cartels: Behind the drug war. CNN (2014, febrero 25)
14. Planas, R. A Murder Every Half Hour in Mexico's Drug War. NYDailyNews.com: Daily News (2012, enero 13)
15. Borderland Beat, http://www.borderlandbeat.com/2011/09/nuevo-laredo-silent-war.html
16. Goodman, J. D. In Mexico, Social Media Become a Battleground in the Drug War. The New York Times: The Lede (2011, septiembre 15)
17. Stevenson, M. Woman Decapitated in Mexico for Web Posting. Associated Press (2011, septiembre 24)
18. Blog del Narco, http://www.blogdelnarco.com/
19. Chaski, C. E. Who's At The Keyboard? Authorship Attribution in Digital Evidence Investigations. International Journal of Digital Evidence. 4(1), 1-13 (2005)
20. Baayen, H., van Halteren, H., Neijt, A., & Tweedie, F. An Experiment in Authorship Attribution, pp. 29-37. En actas de JADT 2002: Sixth International Conference on Textual Data Statistical Analysis (2002)

21. Spassova, M. S. El Potencial Discriminatorio de las Secuencias de Categorías Gramaticales en la Atribución Forense de Autoría de Textos en Español. Tesis doctoral, Universitat Pompeu Fabra, Barcelona (2009)
22. Stamatatos, E., Fakotakis, N., & Kokkinakis, G. Computer-Based Authorship Attribution without Lexical Measures. Computers and the Humanities. 35, 193-214 (2001)
23. Burrows, J. Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship. Literary and Linguistic Computing. 17(3), 267-86 (2002)
24. Corney, M. Analysing E-mail Text Authorship for Forensic Purposes. Tesis de maestría, Queensland University of Technology (2003)
25. Orebaugh, A., & Allnutt, J. Classification of Instant Messaging: Communications for Forensics Analysis. The International Journal of Forensic Computer Science. 4(1), 22-28 (2009)
26. Tambouratzis, G., & Vassiliou, M. Employing Thematic Variables for Enhancing Classification Accuracy within Author Discrimination Experiments. Literary and Linguistic Computing. 22(2), 207-224 (2007)
27. Hall, M. A. Correlation-based Feature Selection for Machine Learning. Tesis de maestría, The University of Waikato (1999)

Hand Vein Infrared Image Segmentation for Biometric Recognition

Ignacio Irving Morales-Montiel[1], J. Arturo Olvera-López[1], Manuel Martín-Ortíz[1], and Eber E. Orozco-Guillén[2]

[1]Facultad de Ciencias de la Computación
Benemérita Universidad Autónoma de Puebla
Av. San Claudio y 14 sur. Ciudad Universitaria.
Puebla, Pue., Mexico
cpycon@hotmail.com,{aolvera,mmartin}@cs.buap.mx
[2]Programa de Ingeniería en Informática
Universidad Politécnica de Sinaloa
Carretera Municipal Libre Mazatlán Higueras Km. 3.
Mazatlán, Sin., Mexico
eorozco@upsin.edu.mx

Abstract. In this paper we propose a technique for both finding vein regions from thermal dorsal hand images and extracting features for biometric recognition; our technique analyzes the geometry of the hand to isolate the vein regions and extract some features (vein bifurcations and ending points) for being used as features in the training sets for classifiers. Commonly, the features extracted are used as geometric and descriptive representation of the vein patterns which are matched with hand vein images in a database in order to determine/verify the person's identity.

Keywords: Biometry, Hand vein recognition, Infrared image segmentation, Dorsal hand vein image processing

1 Introduction

It is very common that for security applications, the users must claim their identify in order for accessing to several services. Typical methods for identification use ID cards, passwords, Personal Identification Numbers (PIN), etc. but they are unreliable since they are easy to be either stolen or forgotten. Currently, it is possible to be aimed by technology through Biometric Recognition systems for overcoming these drawbacks.

In Computer Science, Biometric Recognition is the science of identifying a person based on the physiological and/or behavioral characteristics; these characteristics are known as biometrics and for every person, they have special features such as: Universality, Uniqueness, Permanence and Measurability [1]. Some examples (Fig. 1) of typical biometrics are: Iris, face, palm print, signature, finger print, ear structure, among others [2, 3, 4].

Biometric Recognition is commonly an automatic technique used in our digital age; in this context, a computer system designed for either person identification or person verification is named biometric system. In person verification, the user claims an identity and the recognition system verifies whether the claim is genuine or not; in person identification the recognition system identifies the user without claiming an identity[1].

A typical biometric systems consist of the following phases:

-Capture. This stage is related to the acquirement of the biometric through a digital capture device or sensor (e.g. camera, scanner, microphone).

-Pre-processing. This module respects to the steps for preparing the information to be analyzed from the biometric. When the biometric is stored via a digital image, typical pre-processing steps are: cleaning (noise), contrast enhancement, sharpen, equalization, among other processes. Pre-processing is an important phase since the better quality of information provided to the biometric system the higher recognition rate is achieved.

-Feature extraction. This stage extracts the most relevant descriptive components (minutiae) that describe to the biometric. The main processes involved in feature extraction are the Region Of Interest (ROI) extraction and segmentation; the goal of both processes is the isolation of the region where there exists relevant and descriptive information about the biometric.

-Recognition. This is the last phase in a Biometric Recognition system and uses the features extracted to either person identification or person verification. Commonly, classifiers (supervised, non supervised) are used in this recognition stage; the main goal of the classifier is to extract (through its learning steps) models or patterns from each biometric in the dataset which are used for recognition purposes.

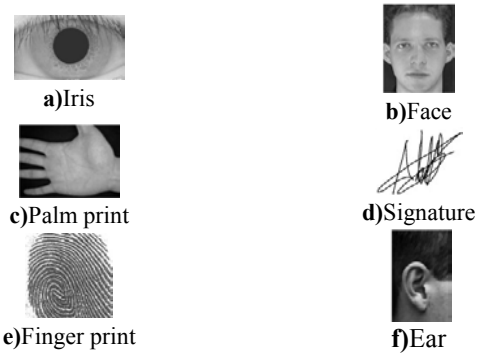


Fig. 1. Biometrics commonly used for person recognition **a)**Iris, **b)**Face, **c)**Palm print, **d)**Signature, **e)**Finger print, **f)**Ear

In this work, we focused in the pre-processing stage, ROI extraction and feature extraction for a Biometric Recognition system based on dorsal hand vein as biometric.

Compared to typical biometrics, vein recognition is a new member of the biometric family. This technology started in the 1990s, but it did not attract much attention in that decade. Since 2000 year, few papers on this topic have been published and nowadays it is an active research field.

Hand vein images are commonly captured by digital cameras operating in the visible range (400nm-700nm wavelength) of the electromagnetic spectrum i.e. illuminating the region to capture using visible light. In the captured images it is difficult to discern the veins since in some cases the person's veins are not superficially visible.

Fortunately, the Infrared (IR) light (2.5-50μm wavelength) could help us when veins are not easily visible. IR light is a non visible light that penetrates into the skin about 3mm. Due to the hemoglobin properties in the veins, it absorbs the infrared radiation and as consequence, the veins are contrasted in a dark color. When IR light is used to capture dorsal vein images, it is necessary to capture via IR cameras which are named thermal cameras and the captured images are named thermal images [5].

One of the main problems using thermal image for recognition is the low contrast in the images, because it is not easy to distinguish components that could be simple to recognize for the human eye. Unfortunately there is not a global solution to this problem, but there are useful methods for specific cases.

It is important to have a special mention about the vein pattern, which is a vast network of blood vessels underneath a person's skin. Like fingerprints, the shape of vascular patterns is believed to be distinct among different people [6-8], and very stable/invariant over a long period of time in most of the cases. In addition, as the blood vessels are hidden underneath the skin and they are mostly invisible to the human eye, vein patterns are more difficult to forge by intruders (in comparison to other biometrics such as signature).

A key part of hand dorsal vein biometric is the difference with other similar biometrics because its recognition requires analyzing alive people, putting special attention that when person is not alive the vein pattern color would change, also the pressure, temperature and other similar characteristics. Hand veins can prevent a possible misuse of users' identity, considering that there are several myths about the misuse with the fingerprints when they do not require vitality by the biometric(e.g. mutilations by intruders that would damage both the users' integrity and users' body).

The properties of hand vein make it a potentially good biometric which offers secure and reliable features for person's identification or verification.

2 Hand vein IR image pre-processing

As it was mentioned before, it is not easy to obtain a pattern from a raw hand vein thermal image. For this reason, the ROI extraction, segmentation and contrast enhancement are necessary for the recognition step in a Biometric Recognition system. In the following sections, we present an approach for ROI extraction, vein segmentation and feature extraction from hand vein thermal images.

2.1 ROI extraction

In Fig. 2, a couple of hand vein thermal images with low contrast are shown. In order to extract the ROI in these low-contrast cases, it is necessary to isolate the hand in the image, so we need to use a mask in order to delimit the hand.

As initial step, we propose to use a 3x3 Median filter to eliminate the possible noise of the image. After that, to get the position of the hand, we use a Sobel Filter to find the area with the most of edges. After edges are obtained, we apply a horizontal white filling (6px) that is a simple process which take every white pixel and look at the 6 neighbor pixels left (in this case), and if it finds a white pixel then it fills the pixels to white until that pixel; then a vertical white filling process is applied, which takes every white pixel and fills in white every pixel above (Fig. 3b).

For each column, the mean is calculated to create a histogram and find the valley points (see Fig. 4a) where:

P1 is originally the valley point between the small finger and the ring finger.

P2 is originally the valley point between the index finger and the thumb.

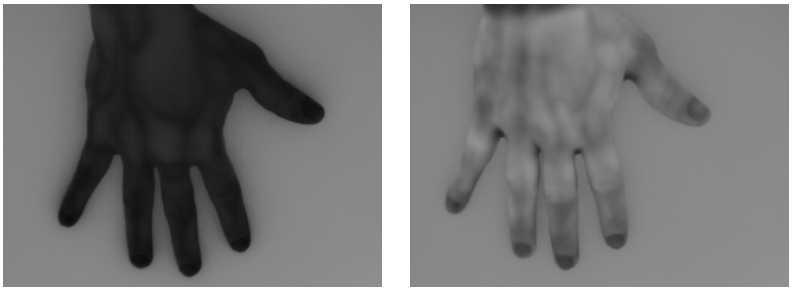


Fig. 2. IR hand vein images of two hands captured in different environments

a) **b)**

Fig. 3. Creating the Mask. **a)**IR hand vein image. **b)** Hand Mask obtained from a)

The Euclidean Distance *D* between *P1* and *P2* is calculated to be used as a reference of the hand dimensions. Then point *P2* is adjusted to estimate the position of the valley point between the middle finger and the index finger (*X* position is *D*4/7* from *P1* and *Y* position is at the same height than *P1-D/7*). *P1* is adjusted to estimate the actual position of the valley point between the middle finger and the index finger (*Y* position is *Y+D/1.4*).

The ROI is therefore defined as a rectangular region $R_{P1P2P3P4}$ (Fig. 4b). Where $\overline{P1P3} = 1.4 * \overline{P1P2}$. The process of defining the ROI is shown in Fig. 3 and Fig. 4.

The reason for extracting the ROI in this way is because it ensures all the ROIs reference to the same region in the hand image, and it is irrelevant to the size of the hands; in addition, extracting ROI based on hand landmark will increase the tolerance of the system against hand rotation and low contrast.

The image depicted in Fig. 4b shows the result of the extracted ROI from the hand image in Fig. 3a.

2.2 Image enhancement and segmentation

The vein patterns show different clearness among different images. That is why it is necessary to enhance the contrast in the images, for this purpose, we propose to apply a 5x5 Median Filter to remove the speckling noise in the images. Then a 2-D Wiener filter as suggested in [9] was applied to the ROI in order to eliminate the high frequency noise. A Wiener filter is useful as lowpass-filter in the grayscale image in order to eliminate noise. Wiener filtering is based on using pixel neighborhoods of size *m*x*n* to estimate the local mean (μl) and local standard deviation (σl). After computing μ_l and σ_l, a filtered pixel $b(i,j)$ is created according to (1) where v^2 is the average of all the local estimated variances. In this paper a 2D-wiener filter is used in 4x4 neighborhood regions over the image $I(x,y)$; the noise is assumed to be additive noise (Gaussian white noise).

$$b(i,j) = \mu_l + \frac{{\sigma_l}^2 - v^2}{{\sigma_l}^2}(I(i,j) - \mu_l) \tag{1}$$

Due luminosity changes in the hand vein images, a global thresholding is not useful for separating the vein pattern from the background i.e. vein segmentation. In order to segment dorsal hand vein images it is necessary to apply an adaptive thresholding algorithm as suggested in [9, 10], which dynamically uses different threshold values for every pixel in the image based on the analysis of its surrounding neighbors.

For every pixel in the image, its threshold value is set as the regional mean value subtracted by a global offset. To determine the local threshold the expressions (2), (3) and (4) are used; where $m(x,y)$ is the local mean, $\partial(x,y)$ is the local deviation, $w>0$ is the size of the area around the pixel and $k \in [0,1]$ is a bias which is a control for the level of adaptation varying threshold value.

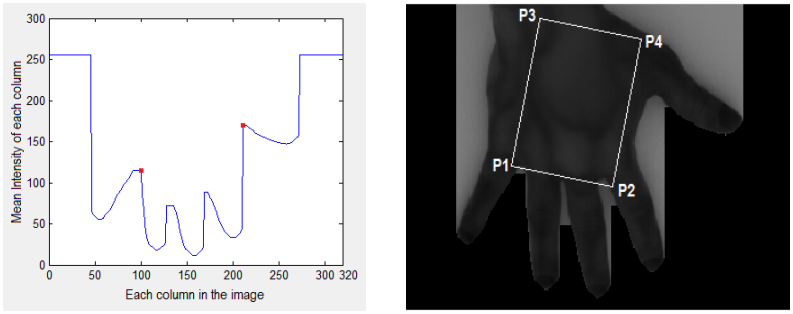
$$m(x,y) = \frac{1}{w^2}\left(\sum_{i=x-w/2}^{x+w/2} \sum_{j=y-w/2}^{y+w/2} I(i,j)\right) \tag{2}$$

$$\partial(x,y) = I(x,y) - m(x,y) \tag{3}$$

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{\partial(x,y)}{1-\partial(x,y)} - 1\right)\right] \tag{4}$$

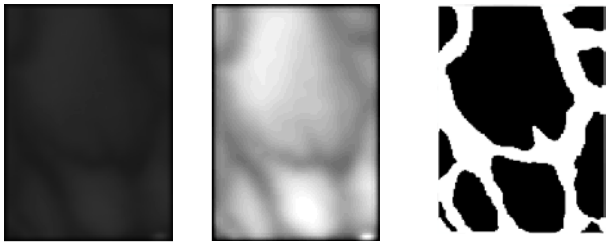
For segmenting the hand image, $T(x,y)$ is used as threshold binarization using w=11.

In Fig. 5b the vein region image after ROI enhancement with noise reduction and normalization is shown. Fig. 5c shows the enhanced ROI after applying the local thresholding binarization.



a) **b)**

Fig. 4. a)Intensity Histogram. **b)** Locating the ROI.



a) **b)** **c)**

Fig. 5. Preprocessing of IR vein pattern images. **a)**ROI for the hand image in Fig. 3a). **b)** After noise reduction and normalization. **c)** After local thresholding binarization

Table 1. Features of the ETIP- 7320 P-Series Infrared Camera Highlight

Detector	Microbolometer 320 x 240 UFPA VOX
Spectral Response	7 to 14 microns
Video Update Rate	50-60Hz (16bit digital)
Accuracy	±1°C or ±1%
Thermal Sensitivity	27 mk
Operating Temperature	-20C to 50C

2.3 Feature Extraction

Once the hand vein images have been segmented it is necessary to extract the main descriptive features (minutiae) for each vein pattern from the image since the features extracted are the baseline in the training sets for the classifiers in the recognition step. As result of the segmentation process, the hand vein images are binary but they can be thinned in order to reduce each vein region to a minimal structure and as result the feature extraction process is applied over vein regions of at most one pixel width. For thinning we used the algorithm based on connectivity skeleton proposed in [11].

The main minutiae points that describe the shape and structure of a hand vein are the bifurcation points (i.e. intersection or cross points between veins) since their amount and position are unique for each person; these bifurcation points are quite similar to delta points in fingerprint images. In this work we propose extracting both bifurcation (as proposed in [12] but applied to hand dorsal veins) and end points of the vein.

According to the characteristics of a bifurcation point, a set of masks is used to find a bifurcation over a vein region; the masks used in this work are shown in Fig. 6, these masks uses as pivot the pixel in the center and matching them, either bifurcation or end points over a vein region are found.

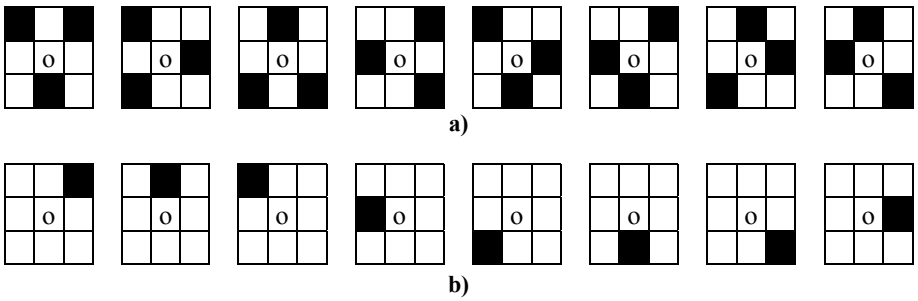


Fig. 6. Masks used for: **a)**Finding bifurcation points. **b)**Finding end points. In both cases the pivot is the center pixel in the 3x3 region

3 Experimental Results

For applying our ROI extraction and segmentation approach, we used IR hand vein images with different levels of contrast. These images were captured via an ETIP 7320 P-Series Infrared Camera described in Table 1.

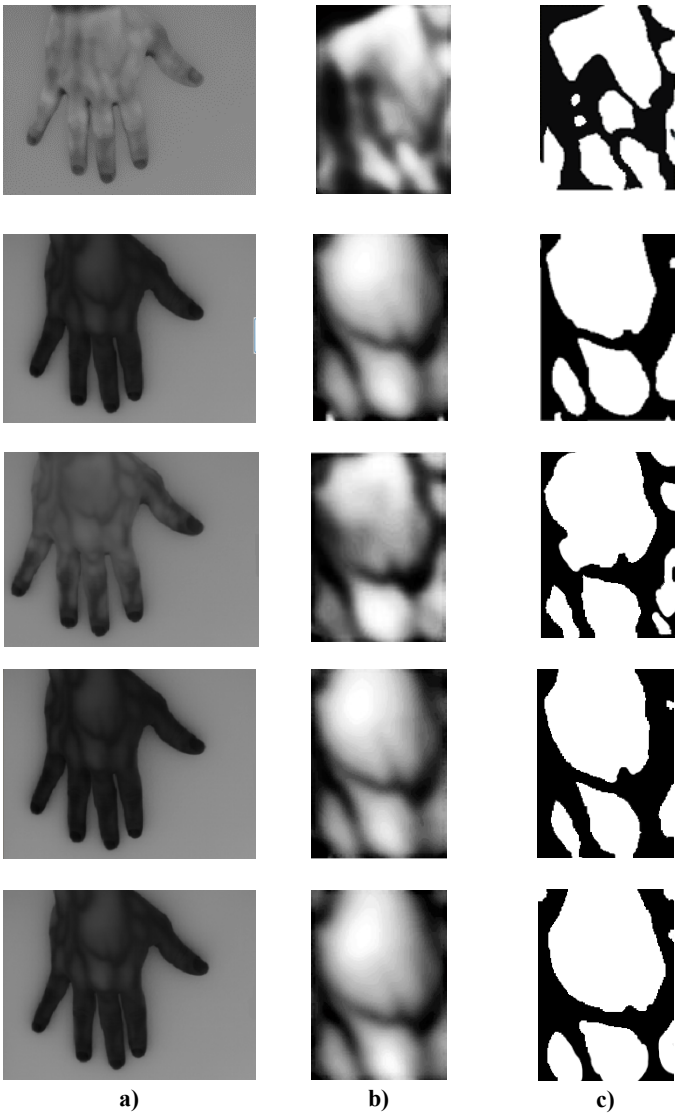
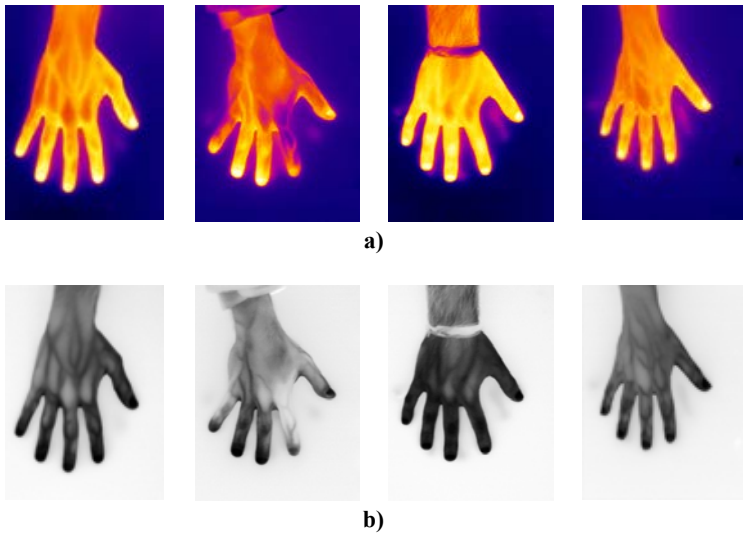


Fig. 7. Results obtained for segmenting IR hand vein images. **a)** Original IR hand vein image. **b)** After ROI extraction, noise reduction an normalization. **c)** After local thresholding

Some results obtained by our ROI extraction and segmentation approach are shown in Fig. 7, where it can be noticed that our approach is able to segment and isolate the ROI related to dorsal hand vein IR images with different low contrast levels which is the main problem to face when trying to analyze biometric hand vein digital images.

A second stage in our experimentation consists in applying our whole approach i.e. ROI extraction, segmentation and feature extraction, as was explained in section 2.

For these experiments we used the *Tecnocampus* thermal hand dataset [13]. In Fig. 8a, some of the dorsal hand thermal images (false thermal color) from this dataset are depicted; for our experiments the images were firstly filtered applying negative and gray scale conversion respectively (Fig. 8b).



a)

b)

Fig. 8. a) Thermal hand datasets obtained from[13]. **b)** Filtered images applying negative and gray scale conversion

Some of the results obtained by applying our whole approach are shown in Fig 9, in this figure the following column images are depicted:

a)Dorsal hand vein thermal image.

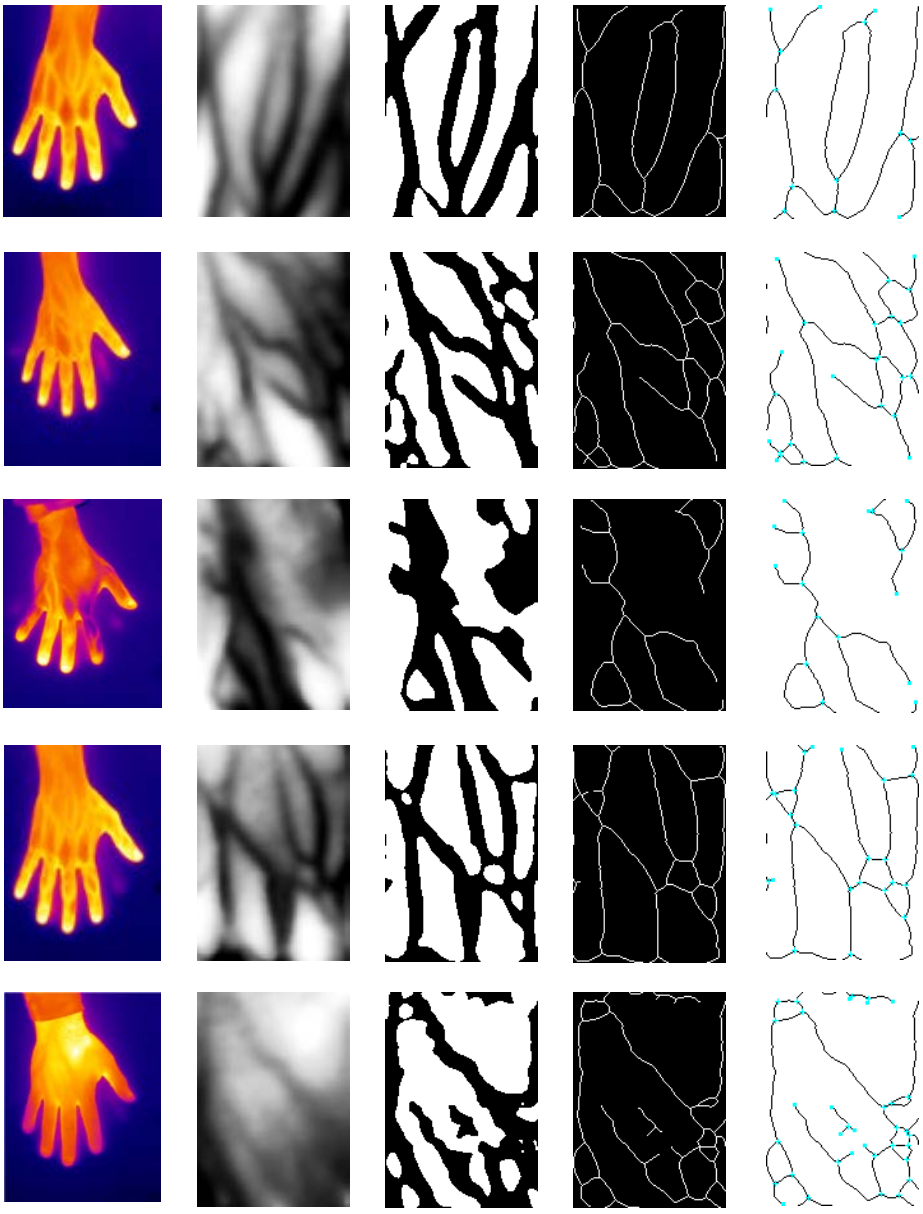
b)ROI extraction, noise reduction and normalization where each hand image is enhanced in order to obtain a better contrast.

c)Hand vein segmentation using local thresholding where each segmented vein region is in black color.

d)Thinning of the vein regions obtained after segmentation.

e)Bifurcation and end points found as minutiae. For a better appreciation of the points, these results are show in negative with respect to the thinning image result.

Based on the results shown in Fig. 9, it can be seen that in most of the cases, the main structures of the hand vein are found and also their respective bifurcation and end points. In addition, the proposed method can be used for images where there exists low contrast in the hand vein image such as the image shown in the bottom of Fig. 9.



a) **b)** **c)** **d)** **e)**

Fig. 9. a) Thermal hand dorsal images. **b)** ROI enhanced and normalized. **c)**Segmentation obtained over images in b). **c)**Hand vein thinning. **d)**Bifurcation and end points found

4 Conclusions

This paper proposes a technique to segment the vein patterns in the human hand, and then extracting minutiae points. This technique may be used to process low contrast images, even when it is not a global solution because it can be affected by motion or rotation; it is enough to be used in several cases. It is very important to highlight that with this approach is possible to extract features (minutiae) from hand vein thermal images with different level of low contrast. Hand vein is considered as biometric because it has a special characteristic since for recognizing a person through this biometric it is necessary that the person stays alive.

This initial research work is related to the pre-processing, ROI extraction and feature extraction stages in a biometric system. Then, as future work we will analyze our approach applying/proposing other thresholding and feature extraction methods in order to construct training sets to be used by classifiers such as Hamming k-Nearest Neighbors, Neural Networks and Support Vector Machines in the recognition step. Finally, an analysis of the recognition results is also included in our future work.

Acknowledgment. This work was supported by the projects OLLJ-ING14-IVIEP-BUAP (Puebla, Mexico) and PROMEP/103.5/12/7954UPSIN-PTC-015 (Sinaloa, Mexico).

References

1. Jain, A. K., Ross, A. A., Nandakumar, K.: Introduction to Biometrics. Springer, (2012)
2. Samaria, F., Harter, A.: Parameterization of a Stochastic Model for Human Face Identification. In: Proceedings of 2nd IEEE Workshop on Applications of Computer Vision, 138-142 (1994)
3. CASIA iris database.: Institute of Automation, Chinese Academy of Sciences. `http://biometrics.idealtest.org`
4. Fierrez, J., Galbally, J., Ortega-Garcia, J., Freire, M. R., Alonso-Fernández, F., Ramos, D., Gracia-Roche, J.: BiosecurID: a multimodal biometric database. Pattern Analysis and Applications, 13(2), 235-246(2010)
5. Wang, L., Leedham, G.: Near and far infrared imaging for vein pattern biometrics. In: Proceedings of the IEEE International Conference on Video and Signal Based Surveillance, p. 52. IEEE Computer Society (2006)
6. Jain, A. K., Bolle, R. M., Pankatni, S.: Biometrics: Personal Identification in Networked Society, Kluwer Academic Publishers, Dordrecht (1999)
7. MacGregor, P., Welford, R.: Veincheck: imaging for security and personnel identification, Adv. Imaging 6(7), 52-56 (1991)
8. Hawkes, P. L., Clayden, D. O.: Veincheck research for automatic identification of people, in: Hand and Fingerprint Seminar at NPL, 230-233 (1993)
9. Wang, L., Leedham, G., Siu-Yeung, C. D.: Minutiae feature analysis for infrared hand vein pattern biometrics. Pattern recognition, 41(3), 920-929 (2008)
10. Romen, S. T., Roy, S., Imocha, S. O., Sinam, T., Singh, K. M.: A New Local Adaptive Thresholding Technique in Binarization. International Journal of Computer Science Issues 8(6) No. 2, 271-277 (2011)

11. Zhang, T. Y., Suen, C. Y.: A fast parallel algorithm for thinning digital patterns. Communications of the ACM, 27(3), 236-239 (1984)
12. Mishra, K. N., Agrawal, P. C., Tripathi, A., Garg, V.: Minutiae Distances and Orientation Fields Based Thumbprint Identification of Identical Twins. International Journal of Image, Graphics and Signal Processing, 5(2), 51-59 (2013)
13. Faundez-Zanuy, M., Mekyska, J., Font-Aragonés, X.: A new hand image database simultaneously acquired in visible, near-infrared and thermal spectrums. Cognitive Computation, 6(2), 230-240 (2014)

Performance Analysis of Retina and DoG Filtering Applied to Face Images for Training Correlation Filters

Everardo Santiago Ramírez[1], José Ángel González Fraga[1], Omar Álvarez Xochihua[1], Everardo Gutierrez López[1], and Sergio Omar Infante Prieto[2]

[1] Facultad de Ciencias, Universidad Autónoma de Baja California, Carretera Transpeninsular Tijuana-Ensenada, Núm. 3917, Colonia Playitas, Ensenada, Baja California, C.P. 22860
{everardo.santiagoramirez,angel_fraga, aomar,everardo.gutierrez}@uabc.edu.mx

[2] Facultad de Ingeniería, Arquitectura y Diseño, Universidad Autónoma de Baja California, Carretera Transpeninsular Tijuana-Ensenada, Núm. 3917, Colonia Playitas, Ensenada, Baja California, C.P. 22860
sinfante@uabc.edu.mx

Abstract. Recognizing facial images with nonhomogeneous illumination is a challenging task. Retina and difference of Gaussians filtering have been applied to facial images in order to remove variations in illumination. This paper presents a study on how these preprocessing operations affect or improve the performance of the correlation filters in a face recognition task. These preprocessing operations were applied to CMU and YaleB facial databases, which containing images with homogeneous and nonhomogeneous illumination, respectively. Results show that these operations improve the performance of correlation filters for recognizing facial images with variable illumination.

Keywords: Retina Filtering, Difference of Gaussians, Correlation Filters, Facial Recognition.

1 Introduction

A facial image captured in an unconstrained environment is exposed to different light sources. These factors cause nonhomogeneous illumination conditions, which degrade the performance of face recognition algorithms that are based in correlation filters. The main problem is related to local shadows that change the appearance of facial features (eyes, nose, mouth and cheeks) and distort their limits (edges). This affects the performance of correlation filters, which use both the form and the content of the object.

Several approaches to the problem of variable illumination in face images have been proposed in the literature. In [1] the authors proposed an algorithm based on DCT (Discrete Cosine Transform) supported by the reduction of brightness gradient, reduction of spectrum components of low-spatial frequency and

spectral characteristics fusion conditioned on the average intensities. Another approach based on DCT was proposed in [2], where the DCT was employed to compensate for illumination variations in the logarithm domain. Because illumination variations lie mainly in the low-frequency band, an appropriate number of DCT coefficients were truncated to minimize variations of the different illumination conditions. In [3], a hybrid approach of PCA (Principal Component Analysis) and correlation filters was presented, while in [4] a low-pass Gaussian filter was used to estimate the illumination on the face image. Although these preprocessing operations improve the global illumination of facial images, the edge of the facial features is not properly recovered.

An approach to address the problem of nonhomogeneous illumination is the retina and difference of Gaussians (DoG) filtering. However, there is no evidence that these preprocessing operations improve the performance of correlation filters. In this paper, we present a study on how the retina and the DoG filtering influence the performance of correlation filters in the facial recognition task. These preprocessing operations were applied to two sets of facial images, one set with homogeneous illumination and the other with nonhomogeneous illumination. The results obtained in this work show that preprocessing improve the performance of the correlation filter for recognizing facial images with nonhomogeneous illumination.

The remainder of the paper is organized as follows. In Section 2, basic concepts of retina filtering, DoG filtering and correlation filters are presented. In Section 3, the employed evaluation methodology and the obtained results are shown. Finally, Section 4 presents the main conclusions of this work.

2 Theoretical Basis

2.1 Retina Filtering

Retina modeling [5] mimics the performance of the human retina, removing the variations in illumination by combining two nonlinear adaptive functions. Let $f(x, y)$ be a normalized facial image. The first nonlinear function is a low-pass filter given by:

$$F_1(x, y) = f(x, y) \geq G_1(x, y) + \frac{\overline{f_{in}}}{E_v}, \tag{1}$$

where $F_1(x, y)$ is the light adaptation factor, the symbol $\geq$ denotes the convolution operation, $\overline{f_{in}}$ is the mean of the input image and $G_1(x, y)$ is a low-pass Gaussian filter, given by:

$$G_1(x, y) = \frac{1}{2\pi\sigma_1^2} exp(\quad \frac{x^2 + y^2}{2\sigma_1^2}). \tag{2}$$

E_v in Eq. 1 influences directly the amount of edge information to be recovered. Small values for E_v are better than large values for retrieving

edge information of facial features, but they only remove a few variations in illumination. Then, $f(x,y)$ is processed according to $F_1(x,y)$:

$$f_{la_1}(x,y) = (max\}f(x,y)| \ + F_1(x,y)) \frac{f(x,y)}{f(x,y) + F_1(x,y)}. \tag{3}$$

The term $(max\}f(x,y)| \ + F_1(x,y))$ is a normalization factor, where $max\}f(x,y)|$ is the maximum intensity value in the input image. The second nonlinear function works similarly. The improved image is given by:

$$f_{la_2}(x,y) = (max\}f_{la_1}(x,y)| \ + F_2(x,y)) \frac{f_{la_1}(x,y)}{f_{la_1}(x,y) + F_2(x,y)}, \tag{4}$$

where:

$$F_2(x,y) = f_{la_1}(x,y) \geq G_2(x,y) + \frac{\overline{f_{la_1}}}{E_v}, \tag{5}$$

and

$$G_2(x,y) = \frac{1}{2\pi\sigma_2^2} exp(\quad \frac{x^2+y^2}{2\sigma_2^2}). \tag{6}$$

2.2 Difference of Gaussians

The image $f_{la_2}(x,y)$ has a uniform texture. In order to provide the filters with a better discrimination capability, an edge enhancement is performed by applying a DoG filter as follows:

$$f_{bip}(x,y) = DoG \geq f_{la_2}(x,y), \tag{7}$$

where DoG is given by:

$$DoG = \frac{1}{2\pi\sigma_{ph}^2} e^{\ \frac{x^2+y^2}{2\sigma_{ph}^2}} \quad \frac{1}{2\pi\sigma_H^2} e^{\ \frac{x^2+y^2}{2\sigma_H^2}}, \tag{8}$$

where the terms σ_{Ph}^2 and σ_H^2 correspond to the standard deviations of the low-pass filters.

2.3 Composite Correlation Filters

Correlation pattern recognition is based on the selection or creation of a reference signal **h**, called correlation filter, and then determining the degree to which the analyzed image $f(x,y)$ resembles the reference signal [6]. Applying a correlation filter to a test image yields a correlation plane $g(x,y)$:

$$g(x,y) = f(x,y) \geq h(x,y) = \cup \ \ ^1\}F(u,v) \times H \ (u,v)| \, , \tag{9}$$

where $F(u,v)$ and $H(u,v)$ are the fourier Transforms (FT) of $f(x,y)$ and **h**, respectively. Since **h** is a vector, it is reshaped to a bidimensional signal denoted

by $H(u,v)$. The symbols $\not\propto$ $\otimes$ and $\cup$ ^{1} are the element-wise multiplication, the conjugate complex and the inverse of the FT, respectively.

Let $f_i(x,y)$ / T be the *ith* training image and $F_i(u,v)$ its FT. The nonlinear filtering [7] applied to $F_i(u,v)$ is given by:

$$F_i^k(u,v) = \underset{\surd}{F_i}(u,v) \overset{k}{\surd} \exp(i\varphi_{F_i(u,v)}). \tag{10}$$

where $0 < k < 1$ is the nonlinearity factor. Let $\boldsymbol{x_j}$ be a column-vector obtained by scanning lexicographically $F_i^k(u,v)$, and the vectors $\boldsymbol{x_j}$ the columns of matrix $X = [\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_N}]$, the nonlinear Synthetic Discriminant Function (SDF) filter is given by [8,7]:

$$h = X^k((X^k)^{+}X^k) \ \ ^{1}\mathbf{u}, \tag{11}$$

where ^{1} denotes the inverse of matrix, $+$ indicates the transpose operation and $\mathbf{u} = [u_1, u_2, ..., u_N]^{+}$ is a vector of size $\underset{\surd}{T}\overset{\surd}{}$ containing the desired values at the origin of the correlation output for each training image.

The Minimum Average Correlation Energy (MACE) filter is based on the principles of the SDF, but is focused to produce a sharp, high peak for authentic images. Let $D = \frac{1}{N*d}\sum_{i=1}^{N}(X_i X_i\,)$ be a matrix where X_i is a diagonal matrix whose elements correspond to $F_i(u,v)$. Consider also the matrix X defined previously, MACE filter is given by [9]:

$$h = D \ \ ^{1}X(X^{+}D \ \ ^{1}X) \ \ ^{1}u. \tag{12}$$

In a similar fashion, the Unconstrained Optimal Trade-off SDF (UOTSDF) filter produces sharp, high peaks, and it is also tolerant to noise. The UOTSDF filter is given by [10]:

$$h = (\alpha D + \ \ \overline{1 \ \ \alpha}C) \ \ ^{1}, \tag{13}$$

where C is a diagonal matrix whose diagonal elements $C(k,k)$ represent the noise power spectral. This filter is optimal to be used when the input images are susceptible to noise at low-light conditions. Thus the UOTSDF filter is an optimal choice as well due to its reduced complexity that alleviates the need to invert a Gram matrix present in the MACE filter.

3 Retina and DoG Filtering for Enhancing Images for Training Correlation Filters

3.1 Setting Up the Experiment

Figure 1 depicts the process for recognizing a facial image. Given an input image, with either homogeneous or nonhomogeneous illumination, it is preprocessed by the retina and DoG filtering described in subsections 2.1 and 2.2. After the original image is preprocessed, the enhanced image is correlated with stored filters. Each correlation output $g_i(x,y)$ is analyzed in order to find a correlation

peak and measure its sharpness by the Peak-to-SideLobe Ratio (PSR) metric. A person is correctly recognized if $PSR \in \tau$, where τ is a recognition threshold defined experimentally. The PSR metric is given by [9]:

$$PSR = \frac{(peak\ value) \quad \mu_{area}}{\sigma_{area}}. \tag{14}$$

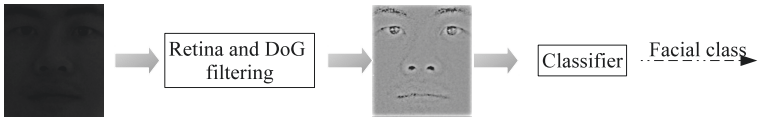


Fig. 1. Process for recognizin an facial image.

The CMU [11] and YaleB[12] facial databases were used to calculate the performance of correlation filters trained with face images preprocessed by retina and DoG filtering. The CMU database contains 13 facial classes. Each facial class consists of a total of 75 different facial images of a same person with homogeneous illumination and different facial expressions. Each image has a resolution of $64 * 64$ pixels. The YaleB database contains 38 face classes. Each class contains between 60 and 64 different facial images of a same person. Each image has a resolution of $192 * 168$ pixels captured under nonhomogeneous illumination.

Fig. 2. Sample of training set for the *ith* facial class in YaleB (row 1) and CMU (row 2).

Let A_j be the set of images of the *jth* facial class. Each set was split into two subsets $T_j \wedge P_j = A_j$, where T_j and P_j are the sets of training images and the

set of test images. Set T_j for the CMU database contains four manually selected face images with different facial expressions, while the set T_j for YaleB database contains four manually selected face images with nonhomogeneous illumination. The test sets P_j have the following characteristics. For the YaleB test set, P_j contains between 56 60 different facial images, while the CMU test set P_j contains 71 different facial images. For evaluation purposes, it is assumed that the images contained in P_j are unknown, so they can be used in order to compute the performance of correlation filters trained with images contained in T_j.

3.2 Numerical Results

Figure 3 shows the process for improving an image with nonhomogeneous illumination. Firstly, the input image $f(x, y)$ is filtered by the low-pass filter given in Eq. 4 in order to produce $f_{la_1}(x, y)$. Secondly, the light adaptation process is performed by applying the low-pass filter given in Eq. 5 to image $f_{la_1}(x, y)$ for producing the image $f_{la_2}(x, y)$. As it can be observed, the local shadows were removed in the image $f_{la_2}(x, y)$. However, the edges of facial features are not easily distinguishable. This problem is solved by applying a DoG filter as is shown in Eq. 7 for obtaining the final image $f_{bip}(x, y)$. This filter produces strong edges for the distinct facial features and illumination variations.

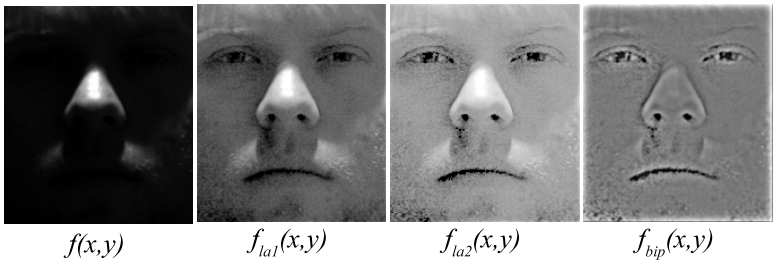


Fig. 3. Enhancing a facial image by appliying retina and DoG filtering. From left to right: input image $f(x, y)$, $f_{la_1}(x, y)$ is the image $f(x, y)$ procesed by the Gaussian filter given in 4, $f_{la_2}(x, y)$ is the image $f_{la_1}(x, y)$ adapted to light acording to Eq. 5, and finally the image $f_{bip}(x, y)$ produced by appliying the DoG filter in Eq. 7 to $f_{la_2}(x, y)$.

Figure 4 shows the PSR performance of a nonlinear SDF filter correlated with 64 different test images of the same class. The PSR values for a filter trained with original images (dotted line) are lower than the PSR values obtained by a nonlinear SDF filter trained with preprocessed face images (solid line). As it can be observed, enhanced images with retina and DoG filtering are easily recognized with a recognition threshold $\tau = 10$.

The FERET protocol states how an evaluation is conducted and how the results are computed in order to measure the performance of face recognition

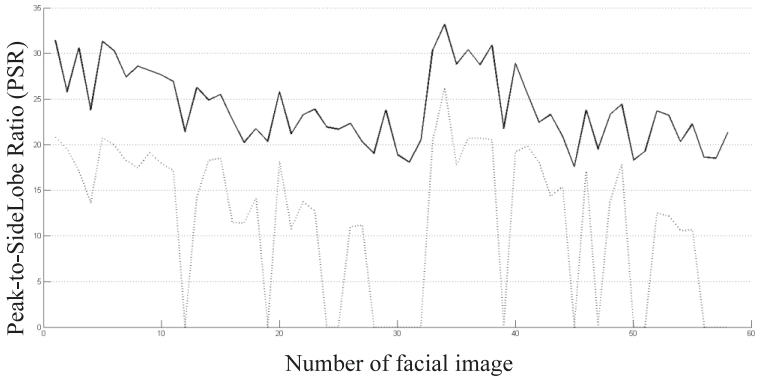


Fig. 4. PSR performance for a nonlinear SDF filter trained with preprocessed face images (solid line) and a nonlinear SDF filter trained with no preprocessed face images (dotted line).

algorithms [13]. Computing performance requires three sets of data. The first is a gallery G which contains correlation filters trained with sets T_j. The other two are probe sets. A probe is a face image $f_j(x, y)$ that is correlated with a filter h_j / G for recognition, where recognition can be verification or identification. The first probe set is P_G which contains different face images of people registered in G (these face images are different from those used to build G). The other probe set is P_N, which contains face images of people who are not in G.

In the identification task, an algorithm determines if a probe $f_j(x, y)$ corresponds to a known person by comparing it against filters $\mathbf{h}_j$ stored in G. If a match is found, then the algorithm identifies the person in the probe $f_j(x, y)$. In the verification task, a person presents his biometric sample (probe $f_j(x, y)$) to a system and claims to be a person registered in G. The algorithm then compares the probe $f_j(x, y)$ with the stored filter $\mathbf{h}_j$ of the person in the gallery. The claimed identity is accepted if there is a match between the image $f_j(x, y)$ and the filter $\mathbf{h}_j$, otherwise, the claimed identity is rejected.

Four FERET metrics were used in this work to compute the performance of the correlation filters. Detection and identification rate (DIR), given in Eq. 15, provides the fraction of probes in P_G that are correctly identified. False alarm rate (FAR), given in Eq. 16, provides the performance when a probe is not of someone in the gallery (i.e., $f_j(x, y)$ / P_N). This type of probe is also referred to as an imposter. The verification rate (VR), given in Eq. 17, provides the performance when a claimed identity is correctly verified. Finally, Eq. 18 shows the false accept rate (FAcR) which provides the times that a claimed identity is declared as authentic, however the identities of the probe $f_j(x, y)$ and the filter h_j) are different.

$$DIR(\tau,1)=\frac{\}f_j(x,y):rank(f_j(x,y))=1\{\ s\ _{,j}\in\tau|}{P_B} \tag{15}$$

$$FAR(\tau)=\frac{\}f_j(x,y):max_i\}s_{i,j}\in\tau|\ |}{P_N} \tag{16}$$

$$VR(\tau)=\frac{\}f_j(x,y):s_{i,j}\in\tau\ \{\ id(\mathbf{h}_j=id(f_j(x,y))|}{P} \tag{17}$$

$$FAcR(\tau)=\frac{\}s_{i,j}:s_{i,j}\in\tau\{\ id(\mathbf{h}_j\not\equiv id(p_j(x,y))|}{(\ P\quad 1)\ B} \tag{18}$$

Function $rank()$ in Eq. 15 sorts in descending order those results where $psr\in\tau$. The identity of a probe $f_j(x,y)$ is associated to the filter $\mathbf{h}_j$ with $rank()=1$. $s\ _{,j}$ is the similarity score obtained by correlating $f_j(x,y)$ with $\mathbf{h}_j$ and both are the same facial class. On the other hand, $s_{i,j}$ is the similarity score obtained when $f_j(x,y)$ is correlated with a filter $\mathbf{h}_j$. In this work, $s\ _{,j}$ and $s_{i,j}$ are computed in terms of the PSR metric.

Table 1. Performance of nonlinear SDF filter using FERET testing protocol.

	CMU				YaleB			
	DIR	FAR	VR	FAcR	DIR	FAR	VR	FAcR
Retina and DoG filtering	98.69	0.55	98.59	3.15	95.53	5.95	95.65	0.44
Original images	98.68	8.18	98.78	0.08	77.76	7.47	77.90	0.03

Table 1 shows the results obtained by the nonlinear SDF filter. Retina filtering improves the performance of the filter in terms of DIR and VR metrics on nonhomogeneous illumination conditions. The values for the FAR metric in both databases has decremented highly using the preprocessed images, while the values for FAcR has increased slightly.

Table 2. Performance of MACE filter using FERET testing protocol.

	CMU				YaleB			
	DIR	FAR	VR	FAcR	DIR	FAR	VR	FAcR
Retina and DoG filtering	98.59	0.00	98.59	0.00	83.53	0.00	83.62	0.11
Original images	94.15	0.00	94.15	0.00	64.50	7.58	64.82	0.20

The performance of MACE filter trained with preprocessed images is improved in all FERET metrics, such as can be observed in Table 2. The UOTSDF filter obtained an improved performance in the YaleB database using images enhanced with retina and DoG filters for recognizing face images with nonhomogeneous illumination. However, the performance remained the same for

the CMU database, which indicates that the retina and DoG filtering does not affect the quality of of face images with homogeneous illumination.

Table 3. Performance of UOTSDF filter using FERET testing protocol.

	CMU				YaleB			
	DIR	FAR	VR	FacR	DIR	FAR	VR	FacR
Retina and DoG filtering	98.37	0.29	98.33	0	87.74	0	87.74	0.01
Original images	98.37	0.29	98.33	0	70.27	4.33	70.27	0.06

The results presented in this section showed that the images enhanced with retina and DoG filtering improve the performance of correlation filters. However, the nonlinear SDF filter obtained the best performance reaching values above 95% in DIR and VR metrics.

4 Conclusions

It has been presented a study of the effect of retina and DoG filtering applied to face images for training correlation filters for face recognition. Applying these preprocessing operations on facial images removes illumination variations significantly, improving the regions affected by local shadows. The edges of the facial features become visible as well as the small details in the facial skin. These changes improve the overall appearance of the facial image, and improve the performance of the correlation filters. The evaluated filters achieved a recognition rate above 94% for images with homogeneous illumination, while for images with nonhomogeneous illumination the filters reached performances greater than or equal to 70%.

Acknowledgments. This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT), with the scholarship number 344833/239152 for author whose name is given first above.

References

1. Forczmanski, P., Kukharev, G., Shchegoleva, N.L.: An algorithm of face recognition under difficult lighting conditions. Przeglad Electrotechniczny (2012) 201–204
2. Chen, W., Joo-Er, M., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Transactions on Systems, Man and Cybernetics **36** (2006) 458–465
3. Savvides, M., Vijaya-Kumar, B., Khosla, P.: corefaces- robust shift invariant pca based correlation filter for illumination tolerant face recognition. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2004)

4. Qu, F., Ren, D., Liu, X., Jing, Z., Yan, L.: A face image illumination quality evaluation method based on gaussian low-pass filter. Proceedings of IEEE CCIS2012 (2012)
5. Vu, N.S., Caplier, A.: Illumination-robust face recognition using retina modeling. In: Proceedings of the 16th IEEE international conference on Image processing. ICIP'09, Piscataway, NJ, USA, IEEE Press (2009) 3253–3256
6. Vijaya-Kumar, B.V.K., Mahalanobis, A., Juday, R.: Correlation pattern recognition. Cambridge University Press (2005)
7. Javidi, B., Wang, W., Zhang, G.: Composite fourier-plane nonlinear filter for distortion-invariant pattern recognition. Society of Photo-Optical Instrumentation Engineers **36** (1997) 2690–2696
8. Casasent, D., Chang, W.T.: Correlation synthetic discriminant functions. Appl. Opt. **25** (1986) 2343–2350
9. Mahalanobis, A., Kumar, B.V.K.V., Casasent, D.: Minimum average correlation energy filters. Appl. Opt. **26** (1987) 3633–3640
10. Kiat-Ng, C.: Pda face recognition system using advanced correlation filters. Master's thesis, Carnegie Mellon University (2005)
11. Lab, A.M.P.: Cmu-face expression database (2001)
12. Georghiades, A., Belhumeur, P., Kriegman, D.: From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Mach. Intelligence **23** (2001) 643–660
13. Phillips, P.J., Grother, P., Micheals, R.: 14. In: Evaluation methods in face recognition. Springer (2005) 329–348

Performance Comparison of Surface Recovering Algorithms used in Fringe Projection Profilometry

Alejandra Serrano-Trujillo, Adriana Nava-Vega

Facultad de Ciencias Químicas e Ingeniería, Universidad Autónoma de Baja California, Tijuana, B.C., México
alejandra.serrano.trujillo@uabc.edu.mx, adriana.nava@uabc.edu.mx

Abstract. Profilometry by fringe projection is an optical metrology technique implemented projecting a sinusoidal grating onto a three-dimensional surface, where the resulting deformed grating image represents the object's height distribution modulated in phase. Fringe projection can be implemented as a phase shifting process or as a one-frame method. In this work, both approaches are compared using a correlation and a Fourier transform based algorithm for phase extraction. As a first step of the surface recovering process, a wrapped phase image is obtained. Therefore, a correction is applied by means of a phase unwrapping algorithm, obtaining a continuous phase distribution. The results show that the phase shifting approach returns a phase distribution with higher resolution and less noise than the one-frame approach.

Keywords: Profilometry, fringe projection, phase wrapping, phase unwrapping

1 Introduction

Profilometry by fringe projection is a non-contact method for measuring the three dimensional surface of an object [1]. It is superior to point-to-point profilometers since it offers a continuous covering of a given surface, it can be carried out much faster, and no mechanical contact is required [2]. Characterizing the roughness of surfaces over an area is an important task for scientific applications as well as for manufacturing processes [3], turning fringe projection profilometry into an important measurement tool in both scientific research and industrial applications. Some examples on the biomedical field include 3D intra-oral dental measurements [4] and lower back deformation measurements [5]; on the industrial field they include corrosion analysis [6], measurement of surface roughness [7] and quality control of printed circuit board manufacturing [8]. Our experimental setup of fringe projection profilometry is shown in Fig 1. It consists of a projector containing a sinusoidal grating, an image acquisition unit and a processing unit. Measurement of shape through fringe projection techniques involves the projection of a structured pattern onto the object surface, the recording of the image of the fringe pattern that is phase modulated by the object height distribution, the process of calculating the phase modulation by analyzing the image using a phase extraction algorithm (involving a phase wrapping and

phase unwrapping process), and finally, calibrating the system for mapping the unwrapped phase distribution to real world 3-D coordinates.

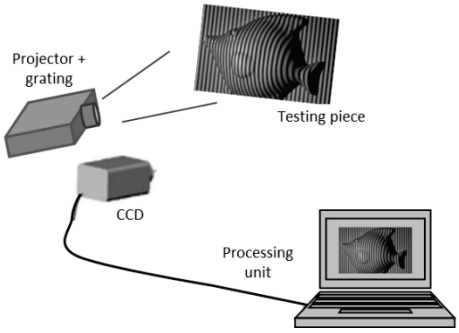


Fig. 1. Fringe projection profilometry setup.

The fringe projection setup is known as a Fourier transform profilometry system if just one fringe projection image is required for surface recovering, otherwise, it is called a Phase shifting profilometry system. The second approach is just a variation of the system described in Fig.1, where the grating is shifted horizontally to cover a period, generating a different fringe projection image to be captured. The grating is then shifted as many times as it takes to complete a fringes period, resulting in a set of images to be processed [9]. Images are then analyzed in order to extract the phase data from them, important because it contains the object's height distribution. This phase extraction process is divided into two steps; a phase wrapping and a phase unwrapping process, described on the following sections.
In this work, both one-frame and phase shifting approaches are compared by means of a Fourier transform [10] and a phase shifting correlation algorithm [11]. In section 2, the two phase wrapping algorithms used for comparison are described. In section 3, the phase unwrapping algorithm is explained. The results are shown on section 4 and finally, conclusions are gathered at section 5.

2 Phase Detection Algorithms

In order to determine an object's height distribution from a fringe projection image it is necessary to find an expression that describes the intensity over the whole image. Phase shifting profilometry is a technique based on the concept of phase shifting interferometry [12]; considered as a part of a set of techniques applied to measure small deformations and displacements in surfaces under test. In interferometry, the fringes appear due an interference of two beams experiment, while in the profilometry technique, the fringes that are projected come from a grating. Given that both systems generate similar fringe patterns, they can be described using the same expression, known as the interferometry equation, defining the intensity in a given coordinate as

$$\mathrm{I}(x,y) = a(x,y) + b(x,y)\cos[\omega x + \varphi(x,y)], \tag{1}$$

where $a(x,y)$ is the average intensity, $b(x,y)$ is the contrast of the fringes, $\varphi(x,y)$ is the phase and ωx is the frequency in the horizontal direction, from the grating. There are a variety of algorithms aimed to detect the phase value based on Eq.(1). Here, a Fourier transform proposal by Takeda [10] and a correlation algorithm originally used for phase shifting interferometry [11] are tested, both proposals are described next.

2.1 Phase Detection by Fourier Transform Algorithm

This method requires only one frame of the deformed fringe pattern to retrieve the surface of the measured object, having an advantage for real time data acquisition and 3-D measurement of dynamic processes. It is also known as Takeda technique. The solution starts from the interferometry equation, rewritten as

$$\mathrm{I}(x,y) = a(x,y) + b(x,y)\cos[2\pi f_0 x + \varphi(x,y)], \tag{2}$$

where $2\pi f_0 = \omega$. Representing Eq. (2) as an exponential expression

$$\mathrm{I}(x,y) = a(x,y) + c(x,y)\exp(2\pi i f_o x) + c^*(x,y)\exp(-2\pi i f_o x), \tag{3}$$

where c^* represents the complex conjugate of $c(x,y) = 1/2b(x,y)e^{i\varphi(x,y)}$.

Applying the Fourier transform respect to *x* to Eq. (3)

$$\mathrm{I}(f,y) = A(f,y) + C(f - f_o, y) + C^*(f + f_o, y). \tag{4}$$

The capital letters denote the Fourier spectra; f is the spatial frequency in the x direction. Any of the two carriers is chosen and filtered. Then, the inverse Fourier transform of $C(f,y)$ respect to f is computed, and $c(x,y)$ is obtained. Finally, the complex logarithm of $c(x,y)$ is calculated as,

$$\log[c(x,y)] = \log[(1/2)\,b(x,y)] + i\varphi(x,y) \tag{5}$$

The phase $\varphi(x,y)$ is then found in the imaginary part of Eq. (5).

2.2 Phase Detection by Means of a Correlation Algorithm

This method is based on the similarity between the captured set of images under a sinusoidal pattern and a real cosine function. The basic equation of phase shifting interferometry is

$$\mathrm{I}_1(x,y) = a(x,y) + b(x,y)\cos[\varphi(x,y) - \alpha_l], \tag{6}$$

where l indicates a shift and α_l is the phase shifting given as $\alpha_l = (2\pi/N)l$. At each shift α_l, a signal $\mathrm{I}_1(x,y)$ is captured, until N shifts have been done. Hence, for each

pixel (x, y) there are N intensity values that should describe a cosine function. A correlation function can then be used to measure how similar is the data respect to the expected cosine form, in order to extract the phase. In the following, the correlation for each pixel $(x\,y)$ of the image will be explored, omitting the notation $(x\,y)$ for brevity. The correlations are analyzed in the z direction that corresponds to the phase step. The correlation function compares two functions $f(z)$ and $g(z)$ as one of them is shifted by an amount ξ respect to the other in the z direction:

$$corr(f,g) = \int_{-\infty}^{\infty} f(z)\,g(z+\xi)dz\ . \tag{7}$$

Then, letting the function $f(z)$ be determined by the observed data points, while $g(z)$ is a simple cosine function, as

$$f(z) = I(z) - \overline{I(z)}, \tag{8}$$

$$g(z+\xi) = aCos(z+\xi), \tag{9}$$

where $I(z)$ represents one of the N captured images, $\overline{I(z)}$ is the average intensity of the N images, $z = -\alpha_l$ and $\xi = (2\pi/m)k$, where k goes from 1 to m, and m is the number of points where the two functions are to be compared. The observed function $f(z)$ should in turn be a function of the unknown phase. According to Eq. (6), $f(z) = a\cos(\varphi + z)$, thus, the correlation of $f(z)$ and $g(z)$ may be written as

$$corr(f,g) = \sum_{l=1}^{N} aCos[\varphi - \alpha_l]Cos\left[\frac{2\pi}{m}k - \alpha_l\right]\Delta, \tag{10}$$

where $\Delta = 2\pi/N$. The correlation attains its maximum when the argument of both cosine functions are equal, that is, when $k = k_{\max}$ such that $\varphi - \alpha_l = (2\pi/m)k_{\max} - \alpha_l$. So the phase is given by

$$\varphi = (2\pi/m)k_{\max}\ . \tag{11}$$

The recovered phase from the deformed fringe pattern by using the aforementioned fringe analysis methods is mathematically limited to the interval [-π,+π] corresponding to the principal value of *arctan* function. Therefore, these routines are known as phase wrapping algorithms. Determination the unknown integral multiple of 2π to be added at each pixel of the wrapped phase map to make it continuous by removing artificial discontinuities is referred to as phase unwrapping [13].

3 Phase Unwrapping by Unweighted Least Squares

The phase unwrapping method applied in this work is the Unweighted Least Squares, and uses the Fast Fourier Transform algorithm [14]. First, the wrapped phase values are extended as a periodic function, performing a mirror reflection over the quadrants. If the wrapped phase matrix is defined in rows and columns as $0 \le i \le M$ and $0 \le j \le$

N, it will end up measuring $0 \leq i < 2M$, $0 \leq j < 2N$. Defining $\psi_{i,j}$ as the wrapped phase function of an original matrix $\phi_{i,j}$, it is possible to obtain a periodic function $\tilde{\psi}_{i,j}$ applying a mirror reflection. In order to find an estimated value of $\varphi_{i,j}$ by means of least squares, the following expressions for phase differences in the horizontal and vertical direction, are computed

$$\Delta^x{}_{i,j} = W\left\{\tilde{\psi}_{i+1,j} - \tilde{\psi}_{i,j}\right\}, \tag{12}$$

$$\Delta^y{}_{i,j} = W\left\{\tilde{\psi}_{i,j+1} - \tilde{\psi}_{i,j}\right\}, \tag{13}$$

where W is a wrapping operator. A function $\tilde{\phi}_{i,j}$ that minimizes the sum of squared differences between the phase differences and those of the solution function is seeked. The least squares solution is the solution of the discrete form of Poisson's equation:

$$\left(\tilde{\phi}_{i+1,j} - 2\tilde{\phi}_{i,j} + \tilde{\phi}_{i-1,j}\right) + \left(\tilde{\phi}_{i,j+1} - 2\tilde{\phi}_{i,j} + \tilde{\phi}_{i,j-1}\right) = \tilde{\rho}_{i,j}, \tag{14}$$

where the extended function $\tilde{\rho}_{i,j}$ and $\tilde{\phi}_{i,j}$ are both periodic, therefore, a Fourier transform can solve the latter expression. Applying two dimensional Fourier transform to the $2M \times 2N$ matrix, on both sides of Eq. (14)

$$\Phi_{m,n} = \frac{P_{m,n}}{2\cos(m\pi / M) + 2\cos(n\pi / N) - 4} \tag{15}$$

where $\Phi_{m,n}$ and $P_{m,n}$ are the two dimensional Fourier transforms of $\tilde{\phi}_{i,j}$ y $\tilde{\rho}_{i,j}$, respectively. The solution of $\tilde{\phi}_{i,j}$ is obtained by applying the inverse Fourier transform to Eq.(15), and the solution $\phi_{i,j}$ is obtained by restricting the result to the dimensions $0 \leq i \leq M$ and $0 \leq j \leq N$.

4 Results

Both correlation and Fourier transform based algorithms were tested using a swordfish bone as a testing piece. Given the length of such piece, the fringe projection technique was applied by dividing the swordfish bone into sections. In the experimental setup the grating used has a sinusoidal fringes pattern of 100 lines per inch. Phase wrapping and unwrapping algorithms were implemented in C language and results were processed using Matlab tools for displaying . Fig. 2 shows the testing piece.

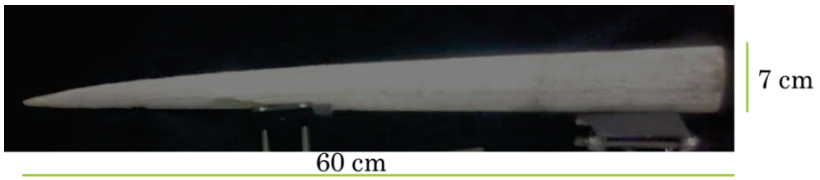


Fig. 2. Testing piece.

Fig.3a shows the complete image of the testing piece, by pasting all the fringe projection sections. Given the high frequency of the grating, the fringes are not distinguished. Fig. 3b shows an extract of 3 sections, where the fringes are visible.

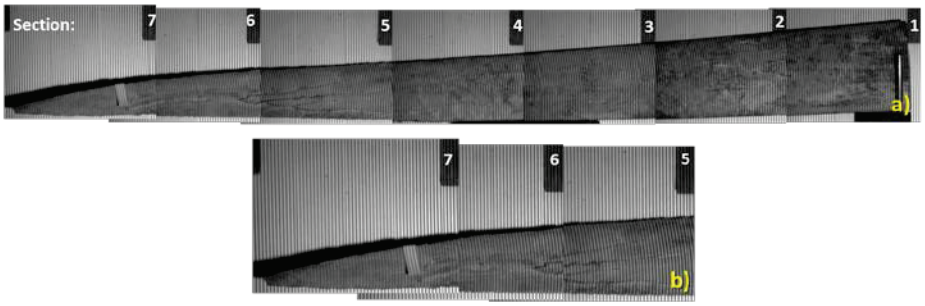


Fig. 3. a) Complete testing piece under fringe projection, b) closer look of sections 5-7

The results at Fig.4a show the wrapped phase image obtained from the Takeda technique, and Fig.4b shows the wrapped phase from the correlation technique, both for only 3 sections of the testing piece. The main difference between them comes from the frequency that the wrapped phase fringes show. This is because the Takeda technique filters out the carrier frequency, which is a high frequency originated from the grating.

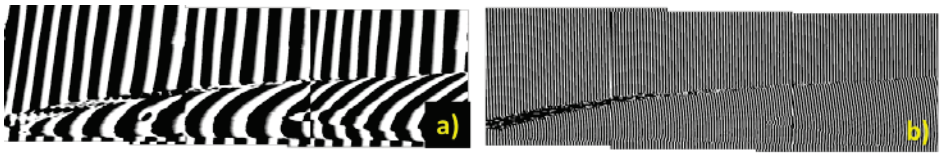


Fig. 4. Wrapped phase: a) by Takeda technique, b) by correlation algorithm.

The performance comparison is based on the phase unwrapping by the unweighted least squares technique. The resulting phase data from the thicker section of the testing piece is shown. Vertical cuts, corresponding to different pixel columns listed on Fig.5 and Fig.6, were plotted in order to demonstrate how well the surface's curvature was detected. Such figures show the differences between the correlation and Takeda algorithms, proving that the correlation algorithm generated a better surface's curvature and texture than the Takeda method in this experiment. In order to prove a better

resolution from the correlation algorithm, the height interval values were computed for every column from the seven sections in which the testing piece was divided. It was found that in average, the correlation algorithm detects a 34.9% more of height values, which means it offers a better resolution. Table 1 shows the range within which the percentage resolution is distributed, representing the extracted phase generated on each section from the swordfish bone.

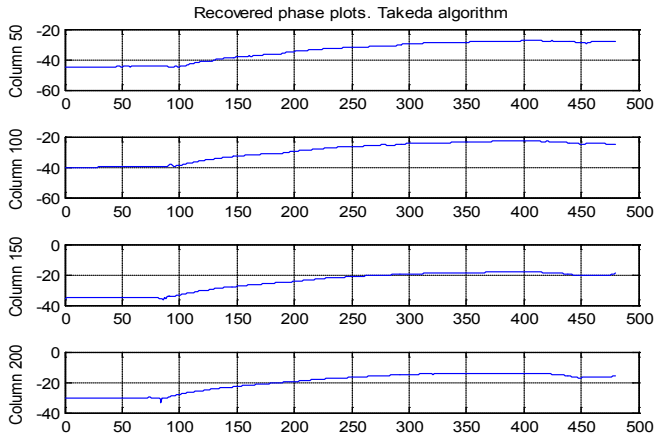


Fig. 5. Unwrapped phase columns from section 1 by Takeda technique

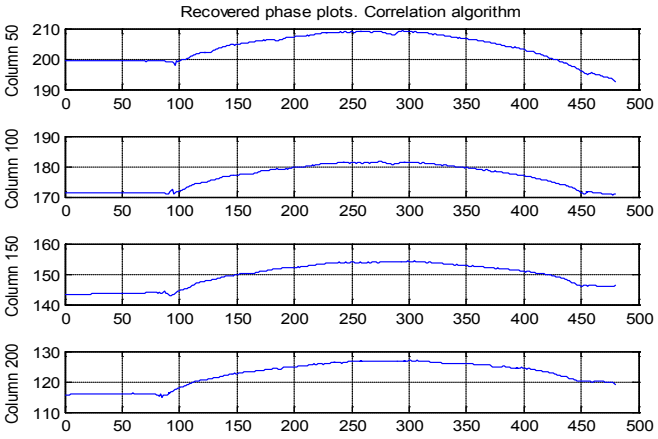


Fig. 6. Unwrapped phase columns from section 1 by correlation algorithm

Table 1. Height resolution from the unwrapped phase of the testing piece.

	Height resolution.				
Testing piece	Takeda alg. (pix)	Correlation alg. (pix)	Δ pix	Normalized to min.	Normalized to max.
Section 1	20.32	33.60	13.28	0.7597	33.2000
Section 2	23.13	26.00	2.87	0.1642	7.1750
Section 3	21.17	25.71	4.54	0.2597	11.3500
Section 4	20.47	25.01	4.54	0.2597	11.3500
Section 5	21.74	22.14	0.4	0.0229	1
Section 6	17.12	19.81	2.69	0.1539	6.7250
Section 7	16.45	33.93	17.48	1	43.7000

Table 1 organizes the resolution (in pixels) obtained for each section of the testing piece according to the approach implemented, then the difference between them is calculated and normalized in the last two columns respect to the maximum and minimum value. From Table 1 it is easy to see that there is a similar performance between the compared approaches for sections 2, 5 and 6. It is seen that section 5 has a similar resolution level for both approaches, while section 7 shows a greater resolution when the correlation algorithm is used. The resolution differences between sections 5 and 7, where normalized values are the unit in Table 1, are mainly related to the occlusions or shadows seen on Fig. 3b, that are interpreted as noise by both approaches. However, since a phase shifting approach works with more than one frame, it presents a lower signal to noise ratio.

Finally, in order to show the difference between these height distributions, Fig. 7 shows the unwrapped phase image of the same section from Figs. 5 and 6. Here, it is easy to see phase unwrapping errors, seen as phase jumps.

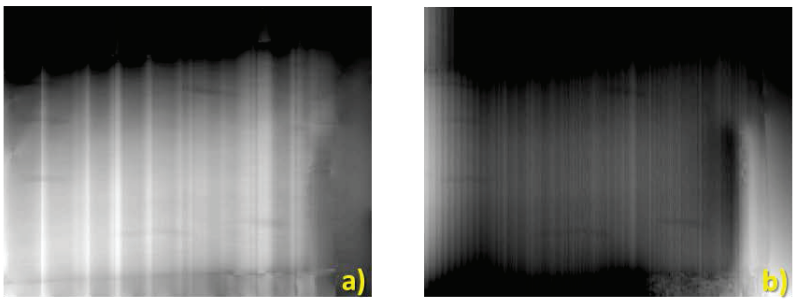


Fig. 7. Unwrapped phase result from section 1: a) by Takeda technique, b) by correlation algorithm

From the experiments done, it was found that the phase shifting profilometry approach using the described correlation algorithm for phase wrapping, and the least

squares algorithm for phase unwrapping, generate an image with phase distribution closer to the real surface than the one-frame approach. Given that a comparison of the real height distribution (mapping from unwrapped phase to 3D coordinates) hasn't been done yet, the results are not presented as a complete quantitative comparison. However, the results obtained from the two approaches implemented; show that for future testing, a phase shifting technique is a better choice. It is important to notice the mount on the right side of Fig. 7b, as well as the horizontal lines that look like scratches over the swordfish bone, since these lines are not evident from the captured images under fringe projection. This means that the algorithm is working correctly based on the frequency of the grating used as well as the image processing algorithms applied, which in the case of Fig.7b show less noise and a better phase distribution.

5 Conclusions

Phase shifting profilometry is a technique that requires more than one fringe projection image in order to recover a surface under test. For such reason, its result presents less noise than a one-frame technique. However, it is important to keep precise grating shifts every time, which can be easily accomplished by using an electrical-mechanical device for controlling the grating position.

This work has compared two fringe projection approaches based on the Takeda and correlation algorithms, the results show that the latter approach brings a better recovery when is post-processed by means of the least squares algorithm. This doesn't mean that the one-frame approach isn't reliable; it depends on several facts related to the image quality and the frequency filtering, in the case of the Takeda method. Depending on the application desired for fringe projection, a one-frame method can be adequate choice, since it offers the advantage of requiring less time and involves an easier capturing process. In this work, the presented phase shifting approach, combined with the correlation phase wrapping algorithm and the least squares unwrapping algorithm is a suitable option for a surface's morphology recovery.

References

1. Malacara D.: Optical Shop Testing, Wiley-Interscience (1992)
2. Salas L., Luna E., Salinas J., Garcia V., Servin M.: Profilometry by fringe projection, Optical Engineering, vol. 42, no. 11, 3307-3314 (2003)
3. Windecker R, Franz S, Tiziani H.: Optical roughness measurements with fringe projection, Applied Optics, vol. 38, no. 13, 2837-2842 (1999)
4. Chen L., Huang C.: Miniaturized 3D surface profilometer using digital fringe projection, Meas. Sci. Techn., vol. 16, no. 5, 1061–1068 (2005)
5. Hanafi A., et. al.: In vivo measurement of lower back deformations with Fourier-transform profilometry, Applied Optics vol. 44, no. 12, 2266–2273 (2005)

6. Huang P. S., Jin F., Chiang F.: Quantitative evaluation of corrosion by a digital fringe projection technique, Opt. Laser Eng., vol. 31, no. 5, 371–380 (1999)
7. Chen L., Chang Y.: High accuracy confocal full-field 3-D surface profilometry for micro lenses using a digital fringe projection strategy, Key Engineering Materials 364-366 (2008)
8. Hui T., Pang G.: Solder paste inspection using region-based defect detection, Int. Journal of Advanced Manufacturing Technology, vol. 42, no. 8, 725–734 (2009)
9. Servin M., Cuevas F.J.: A Novel Technique for Spatial Phase-shifting Interferometry, Journal of Modern Optics, vol. 42, no. 9, 1853-1862 (1995)
10. Takeda M., Mutoh K.: Fourier transform profilometry for the automatic measurement of 3-d object shapes, Appl. Opt., vol. 22, 3977-3982 (1983)
11. Nava-Vega, Salas L., Luna E., Cornejo-Rodríguez A.: Correlation algorithm to recover the phase of a test surface using phase-shifting interferometry, Opt. Express, vol. 12, 5296-5306 (2004)
12. Liang-Chia Chena et al.: 3-D surface profilometry using simultaneous phase-shifting interferometry, Optics Communications, vol. 283, 3376–3382 (2010)
13. Gorthi S. S., Rastogi P.: Fringe projection techniques: Whither we are?, Optics and Lasers in Engineering, vol. 48, 133-140 (2010)
14. Ghiglia Dennis, Pritt Mark: Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software, WileyInterscience (1998)

Real-time 3D Video Processing Using Multi-stream GPU Parallel Computing

Kenia Picos, Víctor H. Díaz-Ramírez, Juan J. Tapia

Instituto Politécnico Nacional, CITEDI
Avenida del Parque 1310, Mesa de Otay, Tijuana, B. C., México
kpicos@citedi.mx, vdiazr@ipn.mx, jtapiaa@ipn.mx

Abstract. This work presents a real-time video processing algorithm for 3D scenes using a graphics processor. The processing is based on parallel computing using concurrent kernels. The proposed algorithm processes individual pixels of each pair of input stereo images to obtain an anaglyph image for each frame. To reduce the computational time, a concurrent kernel implementation using POSIX threads and CUDA streams is utilized. Also, an asynchronous streams execution is used to increase overlapping of the video processing implementation. The obtained results are presented and discussed in terms of speedup and execution time.

Keywords: graphics processing units, parallel computing, CUDA streams, image and video processing.

1 Introduction

Nowadays, the use of parallel computing to solve image and video processing issues in real-time has been rapidly growing [1,7]. Due to the increasing popularity of modern processor designs based on hierarchical memory structure, the use of graphics processing units (GPU) has increased significantly due to their great performance and runtime efficiency [13]. In the state-of-art of image and video processing, there are several algorithms that has been implemented in a parallel platform with an unified device architecture [14]; which results in an increase of the execution performance of the algorithm. In recent applications, a sequential synchronization of the loop iterations within the algorithm is not enough to take full advantage of the benefits of the parallel architecture, by obtaining only a little performance increase with respect to a sequential execution. This is because there are some time gaps in the use of the processing cores, due to synchronization issues. To overcome this shortcoming, an asynchronous algorithm can be used [5]. In this scenario, data communication and the effective use of processing cores are executed asynchronously, avoiding unnecessary waiting times due synchronization, and allowing overlapping in the execution of scheduled tasks [2]. In recent years, a considerable progress has been made within the context of 3D video processing applications. Nowadays, 3D video is one of the latest innovations in information technology [10]. Current

hardware/software technology uses stereo video data to monitor functions as anaglyph display, disparity depth maps computation, and high quality signal processing. In a framework of computer vision technologies, the extraction of 3D object information from 2D images can be carried out by employing modern 3D techniques, such as shape from shading, motion tracking, stereo vision, and many others. Applications such as 3D television and cinema consist on a pair of 2D image sequences which are processed to give us a 3D perception of a scene [9]. A special monitor or glasses are needed to recreate this perception. A 3D video data can be easily generated with a stereo camera, so every observation is interactively changed. However, there are basic limitations that constrain the viewing direction, produce errors when the object presents pose changes, and the execution of the algorithm is inefficient. The last issue can be alliviated by using streams in a GPU implementation.

In the present work, a multi-stream GPU implementation for real-time video processing is proposed. A real-time 3D video processing algorithm is implemented on a GPU with parallel computing using multiple streams. We are focused on improving the performance of 3D video processing, particularly by taking advantage of the asynchronous memory of the GPU. Results obtained with computer simulations using a graphics processor are presented and discussed in terms of computational throughput, and runtime efficiency. The paper is organized as follows. A brief study of multi-stream computation is presented in section 2. The proposed real-time 3D video processing implementation is described in section 3. In section 4, results obtained with the proposed approach are presented and discussed. Finally, section 5 summarizes our conclusions.

2 Multi-stream GPU Computation

The Compute Unified Device Architecture (CUDA) is applied in hardware and software for managing operations in a GPU as a data parallel computing device using an extension to C programming language. This architecture was designed to exploit massive parallelism in graphics processors, and to use concurrent streams for single or multiple GPUs, asynchronous data transfers, and simultaneous kernel executions on a single device. CUDA creates `stream0` by default with the execution of one GPU [11]. Kernel invocations and data transfers are lined up on a stream and processed sequentially in the order they were queued [15]. The creation of multiple streams of execution can perform more work per unit time and make applications to run faster. The tasks are queued on each device (GPU), which can potentially increase the application performance by the number of GPUs installed in the system. Kernel executions can overlap data transfers and data computation to yield real-time processing and reduce the overal runtime on one or more GPUs. In order to improve efficiency on a single device, concurrent kernel executions can be carried out using multiple streams [11]. In contrast to synchronous kernel executions, asynchronous kernel executions are not coordinated within the processor, in where the transmission of messages are allowed to be performed in an unspecified order [4]. Miellou and

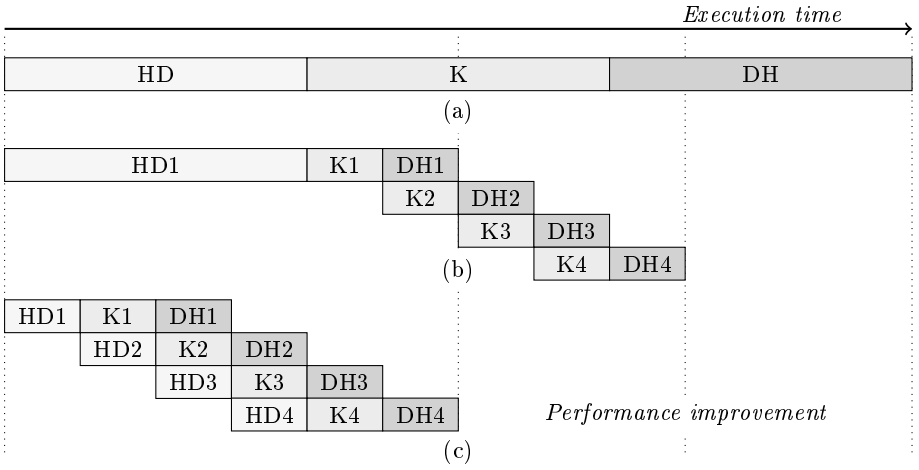


Fig. 1. Concurrent execution of an algorithm using streams. (a) Sequential version. (b) Concurrent version for device-to-host asynchronous memory copy. (c) Concurrent version for host-to-device and device-to-host asynchronous memory copy

Baudet [5] formalized asynchronous algorithms, and described a generalization of an approximation method which includes an inherent parallelism. The main advantages of asynchronous executions are: a) bottleneck reduction, b) synchronization penalty reduction, and c) convergence improvement [5].

3 Real-time 3D Video Processing Using Streams

Concurrent applications can be supported from device by creating a CUDA context for each GPU used in the system. This context can be determined by `cudaStream_t stream[Ns]`, where N_s is the total number of streams. This context includes driver information and capability, such as streams, events, virtual address space, and blocks of memory [12]. With the intercommunication of the context, the device can handle multiple tasks and multiple GPU applications in order to expand parallelism [8]. Several image and video processing applications exploit concurrency through streams. A contribution of this work is the use of a multi-stream approach for implementing stereo disparity estimation and anaglyph video frame generation on GPU in a more efficient way. A sequence of commands that are executed in a sequential order are called streams. These streams are often called by different host threads or POSIX threads (pthreads). The main feature of streams, is that commands can be executed in any order or concurrently. In the present work, we divide the input image in several fragments, and process each fragment asynchronously in a different CUDA stream. This asynchronous behavior allows overlapping of task executions on the device. Fig. 1 shows a stream creation process. A stream is specified by a param-

eter that defines a sequence of kernel invocations including a host-to-device (HD) and device-to-host (DH) memory copy using `cudaMemcpyAsync()` function [11]. The creation of streams could be described with a specific CUDA expression `cudaStreamCreate(&stream[i])`, where i is the number of the current stream. These streams are allocated in a host array in the page-locked memory with `cudaMallocHost(&array, size)`. A kernel call queues the function invocation on the stream related with the current device. In general, a CUDA context for multiple stream execution [12] could be configured by the expression `kernel<<<blocks, threads, 0, stream[i]>>>(parameters)`. For concurrent executions within a device multiple streams are required, that is, multiple kernels can be executed concurrently on a single device [3].
The proposed implementation has been developed in a CPU/GPU architecture running on Linux OS with multi-core host processor (Intel i7) and an NVIDIA graphics processor GeForce GTX780 with 3.5 compute capability. An algorithm with pthreads and CUDA streams for 3D video processing is implemented. The main steps related to the present work consists on a) video capture, b) video processing, and c) video display. In Algorithm 1, we introduce a CUDA/OpenCV interoperability for video capture and display. Pthreads and streams creation is specified in Algorithm 2. Here, memory is copied asynchronously from host to device in order to compute the kernel that is described in Algorithm 3. Then, the memory is copied back from device to host. The anaglyph video frame is generated and displayed in the host using OpenCV framework.

Algorithm 1 Video Capture Algorithm

```
1: procedure VideoCapture(frame)                    ▷ input *.mp4 3D video file
2:     Capturing 3D video frames
3:     fps = capture properties
4:     Memory allocation in GPU
5:     while frame ≠ 0, do                          ▷ video capture is opened
6:         Separate L(x, y) and R(x, y) from frame
7:         L_b = L_g = 0
8:         R_r = 0
9:         Image=VideoProcessing(L, R);             ▷ processing kernel in gpu
10:    end while
11:    FreeCudaMem();                               ▷ free memory allocation
12: end procedure
```

4 Results

In this section, results obtained with the proposed algorithm for 3D video processing are presented and discussed in terms of computational performance and throughput. The algorithm was implemented using parallel computing with pthreads and CUDA streams. In each frame, a pthread executes several CUDA

Algorithm 2 Video Processing Algorithm

```
1: procedure VideoProcessing(L,R)                    ▷ input L, R video channels
2:     Pthread Create and Execute
3:     for i ← 1, Nd, do                             ▷ Nd=number of devices
4:         for i ← 1, Ns, do                         ▷ Ns=number of streams
5:             cudaStreamCreate(&stream[i]);
6:         end for
7:         for i ← 1, Ns, do
8:             j = (i * S * Nd) + (tid * S);         ▷ j is an offset
9:             cudaMemcpyAsync(dL+j, L+j, size, host→device, stream[i]);
10:            cudaMemcpyAsync(dR+j, R+j, size, host→device, stream[i]);
11:            kernel<<<grid,thread,stream[i]>>>(dL+j, dR+j, dI+j);
12:            cudaMemcpyAsync(I+j, dI+j, size, device→host, stream[i]);
13:        end for                                   ▷ Stream destroy
14:    end for                                       ▷ Pthread Join
15: end procedure
```

streams. The maximum number of streams executed per thread depends on the graphics processor architecture. A sample of video frames used in our experiments are illustrated in Fig. 2. Note that CUDA kernel has been used for processing input images, in order to obtain an anaglyph image for each frame. Fig. 2(a) and 2(b) shows the left and right channels of the input sequence. Fig. 2(c) shows the resultant anaglyph images obtained with the GPU. In the anaglyph images, different disparities are computed to present the displayed content. It can be seen that the red-cyan disparity depends on the distance of objects with respect of both channels. The disparity d is obtained from the sum of absolute differences (SAD), as follows:

$$SAD(x,y,d) = \sum_{i,j \in W(x,y)}^{N} |L(i,j) - R(i-d,j)|. \tag{1}$$

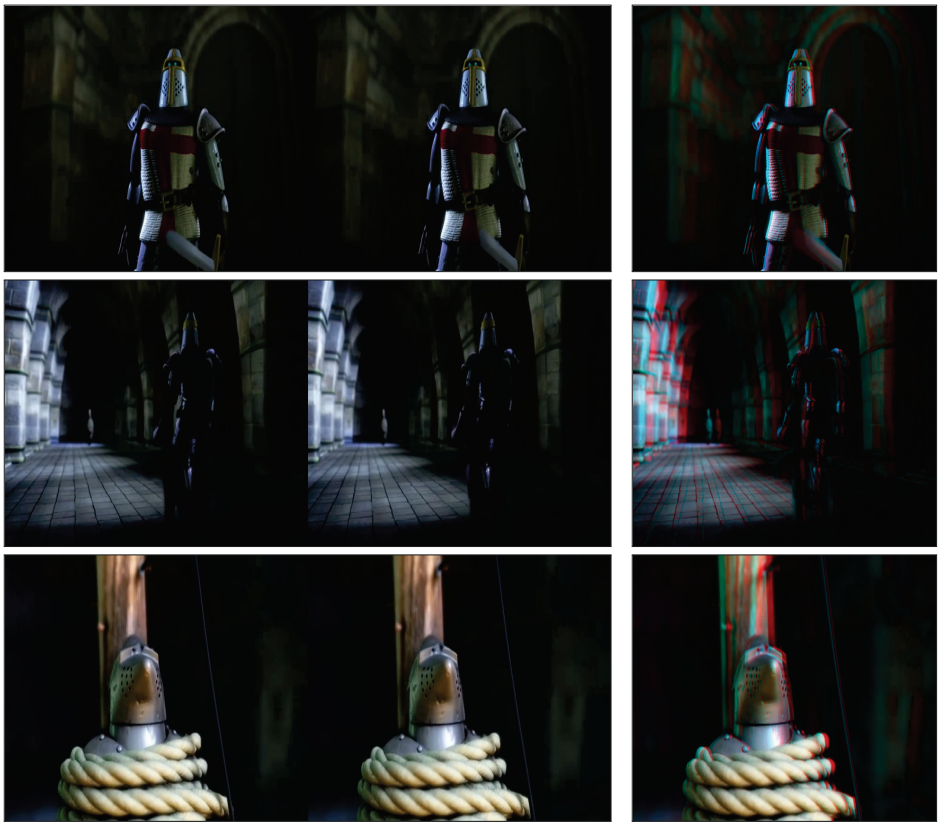
The intensity difference of both channels L and R is calculated for each pixel in a square window $W(x,y)$ with origin at (i,j)-th pixel [6]. The area-based disparity is obtained from the sum of squared differences (SSD) of both channels, given by [6]

$$SSD(x,y,d) = \sum_{i,j \in W(x,y)}^{N} |L(i,j) - R(i-d,j)|^2. \tag{2}$$

The accuracy of the disparity map depends on the window size because there is a correspondence with the probability of matched points. In our experiments the size of the window where W is 11×11 pixels. Note that the size of the window has a direct impact on the computational load of the proposed method. The kernel function used in Algorithm 3 is implemented with 1 to 8 pthreads and with 1 to 8 streams in each pthread. The input images are fragmented in rows and they are processed accordingly with the current thread and stream execution.

Algorithm 3 Kernel function in GPU

1: **procedure** KERNEL<<<GRIDS, THREADS, STREAM[I]>>>(L, R, Image)
2: $x = \text{threadIdx}.x + \text{blockIdx}.x * \text{blockDim}.x$;
3: $y = \text{threadIdx}.y + \text{blockIdx}.y * \text{blockDim}.y$;
4: $tid = x + y * \text{blockDim}.x * \text{grimDim}.x$;
5: $Image[tid] = (L[tid] + R[tid])$;
6: **end procedure**



(a) Left Channel (b) Right Channel (c) Result Image

Fig. 2. Left (a) and right (b) channels. (c) GPU result of anaglyph image at different frames of the sample video sequence (Video sequence from © RedStar Studios)
3D red-cyan glasses are recommended

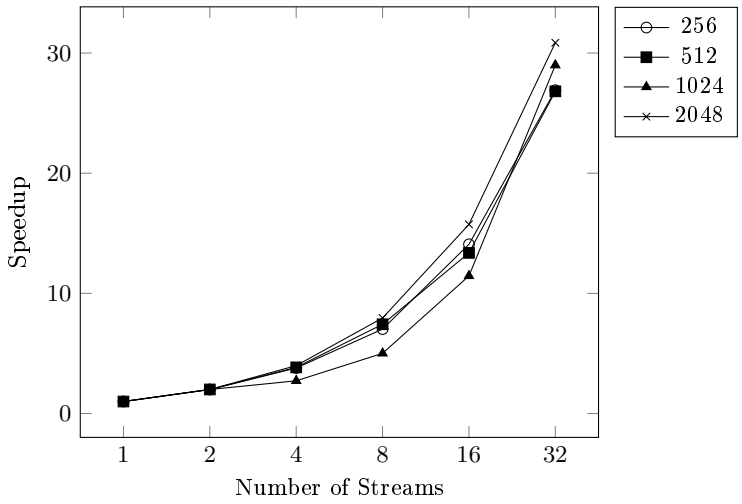


Fig. 3. Performance of the proposed algorithm for different image sizes

The number of grids and threads per block executed concurrently in the kernel depend on the number of streams and pthreads, given by

$$S = (N_x \times N_y) \times CH/(N_d \times N_s). \tag{3}$$

where S is the size of the fragment of the image processed on the current kernel. For each frame, the image fragment has the same size S, and no fragment overlaps with another. $N_x \times N_y$ and CH are the size of the input image and number of channels (3 RGB color channels), respectively. The number of pthreads and the number of streams are represented by N_d and N_s, respectively. The index of a fragment is assigned according the order of call of POSIX and stream argument; however the frames are asynchronously processed. The performance of the algorithm given in terms of the number of threads and streams used in the image partition, is shown in Table 1. It can be seen that by increasing the number of streams per pthread, the speedup increases. The runtime of the kernel execution is 70% off the overall processing executions. The achieved occupancy is the ratio of active warps to the maximum active warps per processor, given by almost 50%. The sample video sequence contains 2000 frames of a set of stereo images with 1024×1024 pixels, and 3 color channels (RGB). Fig. 3 shows the relation of the performance in terms of speedup and number of streams. Observe that with different image sizes, the performance increases more than 30 times when more streams are used.

5 Conclusions

In this work, a CUDA stream-based algorithm that increases parallelism for 3D video processing was proposed. The algorithm can be efficiently used for

Table 1. Computing results for a sample 3D video with 1024×1024 pixels

Threads	Streams per threads	Profiler Time (%)	Kernel execution (ms)	Speedup (times)	Achieved occupancy (%)	Memory stored Throughput (MB/s)	Memory loaded Throughput (GB/s)
1	1	72.63%	15.72	1.00	47.86%	883.64	211.50
1	2	72.48%	7.86	2.00	47.95%	874.76	209.37
1	4	72.49%	3.93	3.99	47.99%	855.03	204.61
1	8	72.32%	1.97	7.98	47.91%	804.08	192.42
2	1	71.90%	7.90	1.99	47.95%	872.57	208.85
2	2	72.19%	5.77	2.72	47.99%	815.72	195.20
2	4	72.51%	3.14	5.01	47.89%	804.19	192.45
2	8	72.98%	1.03	15.31	47.22%	781.47	187.01
4	1	71.06%	3.98	3.95	48.00%	820.25	196.28
4	2	71.83%	2.02	7.79	47.06%	803.98	192.40
4	4	72.80%	1.37	11.45	47.57%	781.66	187.05
4	8	74.33%	0.54	29.22	46.73%	748.43	179.10
8	1	70.29%	2.02	7.77	47.90%	804.09	192.42
8	2	71.69%	1.03	15.20	47.57%	781.45	187.00
8	4	73.68%	0.54	28.99	46.73%	784.60	179.14

real-time 3D video processing. In our implementation, several CUDA streams are executed asynchronously in where each one is able to process a fragment of the input image to compute an anaglyph image. According to computer simulation results, the proposed system yields high effectiveness in the execution of the video processing algorithm in terms on computing performance metrics. The proposed system was implemented in a GPU by taking advantage of massive parallelism. In our implementations, the proposed system achieves a processing rate above 100 frames-per-second (fps) in 3 color channel images of 1024×1024 pixels. The proposed approach performs well with occlusions, however may introduce artificial borders artifacts for a video sequence at speed higher than 70 fps. For future work, the proposed algorithm will be optimized for scenes when the visualization of the object of interest presents light reflections.

References

1. Bhatti, A., Nahavandi, S.: Stereo correspondence estimation using multiwavelets scale-space representation-based multiresolution analysis. Cybern. Syst. 39(6), 641–665 (2008)
2. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms. The MIT Press, 3rd edn. (2009)
3. Farber, R.: CUDA Application Design and Development. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edn. (2012)
4. Hwu, W.W.: GPU Computing Gems Emerald Edition. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edn. (2011)

5. J. M. Bahi, S. Contassot-Vivier, R.C.: Parallel Iterative Algorithms. Chapman and Hall/CRC (2008)
6. Kamencay, P., Breznan, M., Jarina, R., Lukac, P., Zachariasova, M.: Improved depth map estimation from stereo images based on hybrid method. Radioengineering 21(1), 70–78 (2012)
7. Kim, C., Zimmer, H., Pritch, Y., Sorkine-Hornung, A., Gross, M.: Scene reconstruction from high spatio-angular resolution light fields. ACM Trans. Graph. 32(4), 73:1–73:12 (2013)
8. Kirk, D.B., Hwu, W.W.: Programming Massively Parallel Processors: A Hands-on Approach. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edn. (2013)
9. Malik, A.S., Choi, T.S., Nisar, H.: Depth Map and 3D Imaging Applications: Algorithms and Technologies. IGI Global, Hershey, PA, USA, 1st edn. (2011)
10. Matsuyama, T., Nobuhara, S., Takai, T., Tung, T.: 3D Video and Its Applications. Springer Publishing Company, Incorporated (2012)
11. NVIDIA: CUDA C Best Practices Guide (2014)
12. NVIDIA: NVIDIA CUDA Programming Guide 6.0 (2014)
13. Pacheco, P.: An Introduction to Parallel Programming. Morgan Kaufmann (2011)
14. Rodríguez-Sánchez, R., Martínez, J.L., Cock, J.D., Fernández-Escribano, G., Pieters, B., Sánchez, J.L., Claver, J., de Walle, R.V.: 3D high definition video coding on a GPU-based heterogeneous system. Computers & Electrical Engineering 39(8), 2623–2637 (2013)
15. Sanders, J., Kandrot, E.: CUDA by Example: An Introduction to General-Purpose GPU Programming. Addison-Wesley Professional (2010)

Usos del smartphone en actividades académicas realizadas por estudiantes de licenciatura del área computacional de la UABC

Sandra Macías-Maldonado y Javier Organista-Sandoval

Instituto de Investigación y Desarrollo Educativo, Universidad Autónoma de Baja California

sandramaciasm@yahoo.com.mx, javor@uabc.edu.mx

Resumen. Se presenta un estudio sobre los usos educativos del *smartphone* realizados por estudiantes de licenciatura del Área Computacional de la Universidad Autónoma de Baja California (UABC). La muestra fue de 1,073 estudiantes de 30 programas de licenciatura, de los cuales, se seleccionó una muestra representativa de 73 estudiantes del Área Computacional (licenciaturas de Ciencias Computacionales, Ingeniería en Computación e Informática). Los resultados sugieren que los principales usos educativos fueron similares tanto en las licenciaturas del Área Computacional como en las otras licenciaturas. Destacan los usos de comunicación e interacción para lograr acuerdos, realizar tareas y para trabajo en equipo. En cuanto al manejo y acceso a la información, destacan las actividades de búsqueda, intercambio y consulta de contenidos educativos. Mientras que los apoyos organizacionales se enfocaron a consulta fecha/hora, manejo de contactos y recordatorios. Se encontraron diferencias en la proporción de uso, destacando el Área Computacional en la mayoría de ellas.

Palabras clave: Tecnología educativa, Uso educativo del *Smartphone*, Aprendizaje móvil.

1 Introducción

En las últimas décadas, la educación ha sufrido importantes cambios propiciados, en gran medida, por el desarrollo de las tecnologías de la información y comunicación (TIC) las cuales han modificado las formas de acceso y difusión de la información y los modos de comunicación e interacción entre los individuos [1]. En el ámbito educativo es posible considerar a las TIC como herramientas de apoyo al aprendizaje, específicamente como instrumentos que permiten representar de diversas maneras la información y así propiciar la reflexión sobre ella [2].

En particular, la tecnología móvil (dispositivos portátiles) –PDA (personal digital assistant), tabletas, computadoras portátiles, teléfonos celulares, teléfonos inteligentes (smartphones), entre otros- han evolucionado a un ritmo acelerado en la última década, por lo que estos dispositivos cada vez tienen más y mejores características, entre las que se pueden mencionar una mayor capacidad de almacenamiento y procesa-

miento, el uso de contenidos multimedia, diversos medios de conexión, conectividad avanzada a Internet, entre las más sobresalientes [3].

La presencia de los dispositivos portátiles en el ámbito educativo ha cambiado las nociones de tiempo, lugar y espacio de aprendizaje. El espacio de aprendizaje ya no se define solo por la clase con horas programadas en el tradicional salón de clases, los estudiantes pueden aprender tanto dentro como fuera de los salones de clases y/o de la escuela. Estos dispositivos portátiles han hecho posible que el aprendizaje sea facilitado por docentes, por compañeros o propiciado por el mismo estudiante [4]. En este sentido, diversos autores afirman que la presencia de los dispositivos portátiles en las universidades ha cambiado la forma que tienen los estudiantes de comportarse, de interactuar con su entorno y de enfrentarse a sus tareas de aprendizaje [5].

Los resultados aquí presentados se orientan a dar respuesta a las siguientes interrogantes: ¿Cuál es la proporción de smartphones en posesión de los estudiantes de licenciatura del área computacional del campus Ensenada de la UABC? ¿Con qué frecuencia utilizan aplicaciones desde su smartphone? ¿Cuáles son las principales actividades educativas que realizan con mediación de estos dispositivos? ¿Existen diferencias en el uso de aplicaciones y la realización de actividades educativas con mediación del smartphone en las licenciaturas del Área Computacional con respecto de las otras licenciaturas?

2 Marco de referencia

La tendencia tecnológica actual va orientada a una creciente expansión del uso de los smartphones en la mayoría de los países, estos dispositivos son cada vez más accesibles y con mayor penetración en el ámbito educativo. Particularmente, en las universidades existe una tendencia a aumentar el nivel de penetración de los smartphones, lo que sugiere un enorme potencial de apoyo pedagógico.

Los smartphones son considerados dispositivos portátiles con capacidades de comunicación inalámbrica, que facilitan el acceso a Internet y ejecutan funciones o actividades de cómputo básicas. Estas características permiten utilizarlos casi en cualquier parte y en cualquier momento como una herramienta de apoyo educativo ya que les permite a los usuarios estar comunicados, acceder a contenidos educativos, así como ejecutar programas de procesadores de textos, base de datos, lector de libros electrónicos, entre otras funciones [3].

Entre una amplia variedad de definiciones, el aprendizaje móvil se define como aquel aprendizaje que puede ocurrir en múltiples contextos, que se apoya en interacciones sociales y cuyos contenidos educativos pueden accederse desde cualquier dispositivo portátil personal [6]. Un aspecto importante en el aprendizaje móvil tiene relación con la propiedad de los dispositivos portátiles -dispositivos personales de aprendizaje-. Estos dispositivos le brindan al estudiante la posibilidad de apropiarse y personalizar sus propias experiencias de aprendizaje, cuando y donde lo desee, y de capturar sus propios momentos de aprendizaje (fotos, videos, notas, entre otros), así como de interactuar con otros compartiendo sus ideas y preguntas [7].

Tras las imprecisiones teóricas respecto al aprendizaje móvil, por lo reciente de su aparición, en la literatura se refieren cuatro criterios que deben considerarse en la formulación de una teoría del aprendizaje móvil. Estos son: i) la necesidad de distinguir lo que es especial sobre el aprendizaje móvil en comparación con otros tipos de aprendizaje, partiendo del supuesto de que los estudiantes están en movimiento; ii) se debe considerar el aprendizaje que se produce fuera del aula, laboratorios, salas de conferencias, etc.; iii) debe estar basado en relatos de actualidad de las prácticas que permiten un aprendizaje exitoso, y iv) debe considerar el uso ubicuo de la tecnología [8]. En este sentido, en el intento por aportar bases teóricas que sustenten al aprendizaje móvil se han desarrollado propuestas con una diversidad de criterios tanto pedagógicos como tecnológicos para intentar caracterizar a esta modalidad de aprendizaje [9,10,11].

Diversos autores afirman que el éxito de la aplicación del aprendizaje móvil en los contextos educativos dependerá en gran medida del diseño instruccional; de cómo se utilizan los dispositivos portátiles; y de la creación y adopción de políticas y procedimientos de uso y administración de los recursos [3], [12]. En este contexto, algunos estudios tanto a nivel internacional como nacional dan cuenta de los beneficios que se pueden obtener cuando se diseñan ambientes de aprendizaje móvil tanto dentro de los salones de clases como fuera de ellos [13,14,15].

Con base en lo anterior, se observa que el potencial de los dispositivos portátiles como apoyo al proceso educativo brinda la oportunidad de aprendizaje casi en cualquier momento y en cualquier lugar, convirtiendo cualquier momento en una oportunidad para el aprendizaje. Sin embargo, falta mucho por explorar para poder caracterizar el potencial de los dispositivos portátiles en el proceso de enseñanza aprendizaje. En este sentido, en el presente artículo se exploran y caracterizan los usos educativos que con mediación del smartphone están realizando los estudiantes de las licenciaturas del Área Computacional, así como las posibles diferencias en el uso de aplicaciones y realización de actividades académicas entre los estudiantes del Área Computacional y los estudiantes de las otras licenciaturas del campus Ensenada de la UABC.

3 Método

Esta investigación fue realizada en el marco del proyecto de investigación "Tipología del uso educativo de dispositivos móviles 3G realizado por estudiantes y docentes de dos unidades de la UABC en Ensenada", financiado por la XV Convocatoria Interna de Investigación de la UABC. Participaron tres estudiantes de posgrado en el mismo.

Las dos unidades de la UABC consideradas fueron: Ensenada y Valle Dorado, en las cuales se ofrecen 30 programas de licenciatura; la población estudiantil total de las dos unidades en el periodo 2011-1 fue de 9,008 estudiantes (http://csege.uabc.mx).

El algoritmo que se utilizó para la estimación del tamaño de la muestra considerado: tamaño de la población (N=9008), nivel de confianza elegido (Z=1.96, 95%), probabilidad del factor a estudiar (P=0.05) y estimación de error máximo (e=0.03), dando como resultado una muestra representativa de n=1,073 estudiantes [16].

Con el propósito de explorar los usos educativos que con mediación del smartphone están realizando los estudiantes de las licenciaturas del Área Computacional (n=73), en el presente estudio se seleccionaron las licenciaturas de Ciencias Computacionales, Ingeniería en Computación e Informática.

Para la investigación se desarrolló una encuesta de uso educativo de los teléfonos celulares en la universidad [17], cuyo propósito fue conocer el uso del smartphone como herramienta de apoyo educativo. Dicha encuesta estuvo integrada por un total de 60 reactivos, agrupados en cuatro dimensiones:

- *Datos generales*. Se solicita al estudiante: nombre, edad, sexo, carrera que estudia, semestre, dominio del idioma inglés, promedio de calificaciones, disposición de Internet, escolaridad de los padres, tiempo de dedicación laboral, técnica preferida para aprender, entre otras.
- *Aspectos tecnológicos*. Se considera: Posesión de teléfono celular, tipo de activación, velocidad de respuesta del teléfono celular, sistema operativo, tipo de teclado, tipo de cámara, conectividad, así como la frecuencia de uso semanal de las aplicaciones.
- *Aspectos del estudiante*. Se incluyen los siguientes aspectos: tiempo de usar el celular, forma en que aprendió a usar el celular, necesidad e importancia de usar el celular en actividades académicas, gusto y motivación a usar el celular, entre otras.
- *Aspectos estudiante-tecnología.* Se recaba información sobre: la estimación del porcentaje de uso educativo y no educativo que se le da al celular, actividades de comunicación con intención educativa, modalidad de comunicación más utilizada, actividades de manejo de información, actividades de organización, principales usos educativos y opinión de las principales ventajas y desventajas de utilizar el celular como apoyo en los estudios.

En el presente estudio las variables de interés fueron: género, edad, autoconcepto en habilidad tecnológica, disponibilidad de Internet en el hogar, posesión de teléfono celular, posesión de smartphone, principales aplicaciones utilizadas y usos educativos realizados con mediación de dispositivos smartphones.

La aplicación de la encuesta se realizó en el primer ciclo escolar de 2011. Se consideraron al azar salones de clases de las distintas carreras de licenciatura del campus universitario. Los estudiantes respondían el cuestionario dentro del salón de clases en un tiempo promedio de 10 minutos. El análisis de datos se efectuó desde una perspectiva cuantitativa con el paquete estadístico SPSS ver. 17. En el presente estudio se obtuvieron descriptivos básicos, entre ellos, distribución de frecuencias, indicadores de tendencia central e indicadores de dispersión. Lo anterior con el propósito de obtener una descripción general de los datos, el nivel central de la información, así como la variabilidad de la misma.

Tabla 1. Características personales de los estudiantes

	Género		Edad
	Hombre	Mujer	
	n (%)	*n* (%)	*Media (D.E.)*[*]
Área computacional	55 (75.3)	18 (24.7)	21.7 (4.2)
Global	509 (47.5)	562 (52.5)	21.3 (3.8)

*Desviación estándar

Tabla 2. Autodefinición en habilidad tecnología del estudiante

	Autoconcepto habilidad tecnológica			
	no familiar	principiante	intermedio	avanzado
	n (%)	*n (%)*	*n (%)*	*n (%)*
Área computacional	1 (1.4)	2 (2.8)	43 (59.7)	26 (36.1)
Global	4 (0.4)	176 (16.5)	728 (68.0)	162 (15.1)

Escala 0-3; 0=no familiar, 1=principiante, 2=intermedio y 3=avanzado.

4 Resultados

En la tabla 1 se muestran las características personales de género y edad de los participantes del Área Computacional y el Global. En el Área Computacional se observa que predominan los hombres con el 75.3%, mientras que en la media Global se observa un balance entre hombres y mujeres. Con respecto a la edad, la media del Área Computacional fue de 21.7, similar a la edad media Global.

En relación al autoconcepto en habilidad tecnológica (ver tabla 2), la cual se refiere a la percepción del estudiante de sus habilidades generales en el uso de la computadora y el manejo del teléfono celular, 95% de los estudiantes del Área Computacional se ubicaron en el nivel intermedio o avanzado, valor que está por encima del Global (83.1%). Llama la atención el nivel avanzado, donde el Área Computacional se encuentra 21% por encima del Global.

En la tabla 3 se presentan los resultados de las variables disponibilidad de Internet en casa, posesión de teléfono celular y dispositivo tipo smartphone. Con respecto a la disponibilidad de Internet y posesión de teléfono celular, llama la atención que el porcentaje del Área Computacional (80.8% y 94.5% respectivamente) está ligeramente por debajo del porcentaje Global (84.2% y 96.1%). En relación al tipo de teléfono celular, se utilizó la variable conectividad avanzada (presencia de Wi-Fi y/o 3G/4G) para determinar si el dispositivo es smartphone. En el Área Computacional, 57.5% reportó tener un dispositivo tipo smartphone, valor por encima del Global (50.2%).

Tabla 3. Disposición de internet en casa, posesión de celular y posesión de *smartphone*

	Disponibilidad de Internet		Posesión de celular		Posesión de *smartphone*	
	SI	NO	SI	NO	SI	NO
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Área Computacional	59 (80.8)	14 (19.2)	69 (94.5)	4 (5.5)	42 (57.5)	31 (42.5)
Global	903 (84.2)	169 (15.8)	1029 (96.1)	42 (3.9)	539 (50.2)	532 (49.8)

Tabla 4. Porcentaje de uso de las aplicaciones y valor de diferencia entre Área Computacional y Global

	Aplicaciones	Área Computacional	Global	Diferencia AC - G*
		%	%	%
De comunicación	De conferencia	37.9	25.3	12.6
	Correo	38.6	30.7	7.9
	Llamada voz	78.2	72.2	6.0
	Redes sociales	56.1	52.9	3.2
	Mensajes texto	83.7	82.4	1.3
De información	Navegador de Internet	63.9	46.8	17.1
	Lector PDF	37.6	20.9	16.7
	Buscador de información	69.7	53.9	15.8
	Editor texto/ Hoja cálculo	31.3	20.9	10.4
	Calculadora	46.3	37.1	9.2
	Diccionario/Traductor	27.2	20.2	7.0
	Reproductor música	78.1	71.4	6.7
	Grabación audio	26.0	21.7	4.3
	Reproductor video	45.4	42.3	3.1
	Grabación video	30.3	31.2	- 0.9
	Manejo fotos	53.7	58.8	- 5.1
De organización	Editor notas	42.2	34.4	7.8
	Calendario/ Agenda	44.4	50.5	- 6.1
	Manejo contactos	57.1	66.2	- 9.1

* AC: Área Computacional; G: Global

Las principales aplicaciones utilizadas desde el smartphone se clasificaron según su uso en tres categorías: comunicación, información y organización. En la tabla 4 se muestra el porcentaje de uso de las aplicaciones y el valor de diferencia entre el Área Computacional y el Global. Las mayores diferencias a favor del Área Computacional se encontraron en la categoria de información, navegador de Internet (17.1%), lector PDF (16.7 %) y buscador (15.8%). Mientras que las diferencias a favor del Global se encontraron principalmente en las aplicaciones de organización, manejo de contactos (9.1%) y calendario y agenda (6.1%). Con respecto a las aplicaciones mayormente

utilizadas, coincide tanto el Área Computacional como el Global, las más utilizadas fueron mensajes de texto, llamadas de voz y reproductor de música.

Asimismo, las principales actividades educativas realizadas desde el smartphone se clasificaron según su uso en: comunicación, información y organización (ver tabla 5). Las mayores diferencias a favor del Área Computacional se encontraron en las actividades de información, en consulta de información (25.2%) y descarga de información (17.8%). Mientras que las diferencias a favor del Global se encontraron solamente en la actividad de pedir ayuda (2.1%). Con respecto a las principales actividades educativas realizadas, coincide tanto el Área Computacional como el Global, fueron consulta fecha/hora y acuerdos y tareas.

Tabla 5. Distribución porcentual de los usos educativos y el valor de diferencia entre Área Computacional y Global

	Actividades educativas	Área Computacional	Global	Diferencia AC - G*
		%	%	%
De comunicación	Acuerdos y tareas	78.6	70.4	8.2
	Informes/aclaraciones	71.4	65.5	5.9
	Trabajo en equipo	59.5	54.0	5.5
	Pedir ayuda	38.1	40.2	- 2.1
De información	Consulta	66.7	41.5	25.2
	Descarga	47.6	29.8	17.8
	Intercambio	54.8	50.7	4.1
	Búsqueda	76.2	52.4	3.8
De organización	Elaboración notas	52.4	38.4	14.0
	Recordatorios	73.8	62.8	11.0
	Consulta fecha/hora	81.0	74.1	6.9
	Manejo de contactos	66.7	61.7	5.0

5 Discusión

Con respecto al género de los participantes se observa que en el área computacional 75.3% de los participantes son hombres, lo cual coincide con las estadísticas a nivel nacional donde a pesar del incremento en la incorporación de las mujeres a la educación superior (ascendió del 17% en 1969 al 50% en el 2000), todavía en la actualidad representan una cuarta parte en aquellas áreas que tradicionalmente se han considerado masculinas como ingeniería y tecnología [19].

Con relación a la habilidad tecnológica de los estudiantes del Área Computacional (licenciaturas de Ciencias Computacionales, Ingeniería en Computación e Informática), se encontró que 36.1% de los estudiantes se ubicaron en el nivel avanzado en sus habilidades generales en cuanto al uso de la computadora y el manejo del teléfono celular, 21% por encima de las otras licenciaturas (Global). La mayoría de los estudiantes del Área Computacional comparten un interés común en tecnología compu-

tacional, a la vez que van adquiriendo conocimientos y desarrollando habilidades durante su formación profesional en la universidad.

Los resultados sugieren que los estudiantes, tanto del Área Computacional como de las otras licenciaturas, se están comunicando con mediación de su smartphone para lograr acuerdos, para pedir informes, hacer aclaraciones y para trabajar en equipo. Estas actividades educativas las están realizando mayormente con las modalidades de mensajes de texto, llamadas de voz y programas de redes sociales. Resultados similares a los encontrados en la presente investigación se han encontrado en diversos estudios, tanto a nivel local, nacional e internacional; en los cuales, las aplicaciones de comunicación son las mayormente utilizadas por los estudiantes universitarios, destacando envío de mensajes y llamadas por voz [20,21,22,23]. La teoría socio-cultural sostiene que la interacción social, la conversación y el dialogo son fundamentales para el aprendizaje. La conversación mediada con estos dispositivos es propicia para una retroalimentación oportuna y personal, tanto para docentes como para estudiantes [11].

En cuanto a los resultados de manejo y acceso a la información con mediación del smartphone, tanto del Área Computacional como de las otras licenciaturas, los estudiantes están realizando actividades de búsqueda, consulta e intercambio de contenidos educativos. Estas actividades educativas generan un ambiente propicio para tener episodios de aprendizaje mediante la interacción y colaboración entre pares.

Asimismo, tanto los estudiantes del Área Computacional como de las otras licenciaturas, están utilizando sus dispositivos para organizar sus actividades educativas mediante la consulta de fecha-hora, manejo de contactos y recordatorios, principalmente. Estas actividades pueden contribuir en diversos aspectos de la educación de los estudiantes, entre ellas, a que el estudiante asista con puntualidad a clases y exámenes, entrega de trabajos y tareas, registro de notas, entre otras.

Como se puede observar, tanto las aplicaciones mayormente utilizadas como las principales actividades educativas fueron similares tanto en las licenciaturas del Área Computacional como en las otras licenciaturas; sin embargo, los resultados sugieren diferencias importantes en la proporción de uso. Con respecto a las diferencias porcentuales del uso de aplicaciones, destaca el Área Computacional en la mayoría de ellas, especialmente en cuanto al navegador de Internet, lector PDF, buscador de información, de conferencia y editor de texto/hoja de cálculo. Estos programas sugieren un amplio potencial para apoyar el proceso de enseñanza aprendizaje.

Asimismo, se encontran diferencias en el porcentaje de estudiantes que realizan las actividades educativas desde el smartphone. El Área Computacional destaca en todas las actividades con excepción de la comunicación para pedir ayuda, encontrandose mayores diferencias en las actividades de consulta y descarga de información, elaboración de notas, recordatorios y la comunicación para realizar acuerdos y tareas.

Los resultados presentados en el marco de la investigación sugieren que los estudiantes del Área Computacional estan incorporando sus dispositivos smartphones como una herramienta de cómputo educativo, ya que sus conocimientos y habilidades en tecnología computacional les brinda la oportunidad de conocer de manera mas amplia estos dispositivos, apropiandose y usandolos con mayor facilidad. Sin embargo, es necesario que docentes, investigadores y autoridades educativas apoyen y pro-

picien las condiciones necesarias para aprovechar el potecial pedagógico con que cuentan estos dispositivos al incorporarlos al proceso de enseñanza aprendizaje.

Es importante señalar que los hallazgos de esta investigación permitieron que actualmente se desarrolle una investigación a mayor escala en torno al uso educativo de los dispositivos portátiles en la zona urbana de Ensenada con financiamiento de la SEP/CONACYT. Se espera disponer a corto plazo de mayor información del uso pedagógico de los dispositivos portátiles en todos los niveles educativos.

Referencias

1. Cantillo, C., Roura, M., Sanchez, A.: Tendencias actuales en el uso de dispositivos móviles en educación. La Educación digital magazine, 147 (2012)
2. Coll, C., Mauri, T., Onrubia, J.: Análisis de los usos reales de las TIC en contextos educativos formales: una aproximación sociocultural. *Revista Electrónica de Investigación Educativa, 10*(1), 1-18 (2008)
3. Cruz, F., López, M.: Una visión general del m-learning y su proceso de adopción en el esquema educativo. Instituto de Ingeniería, UABC (2007). Disponible en: http://azul.iing.mxl.uabc.mx/~renecruz/papers/ Paper2-Cruz-Flores.pdf
4. Pegrum, M., Oakley, G., Faulkner, R.: Schools going mobile: A study of the adoption of mobile handheld technologies in Western Australian independent schools. Australasian Journal of Educational Technology, 29(1), 66-81 (2013)
5. López, F.A., Silva, M.M.: Patrones de m-learning en el aula virtual. Aplicaciones para el aprendizaje móvil en educación superior [monográfico]. Revista de Universidad y Sociedad del Conocimiento (RUSC), 11(1), 208-221 (2014)
6. Crompton, H.: A diachronic overview of technology contributing to mobile learning: A shift towards student-centred pedagogies. En Ally, M. & Tsinakos A. (Eds.), Increasing Access through Mobile Learning, pp. 7-16. Vancouver: Commonwealth of Learning & Athabasca University (2014)
7. Parsons, D.: The future of mobile learning and implications for education and training. En Ally, M. & Tsinakos A. (Eds.), Increasing Access through Mobile Learning pp. 217-229. Vancouver: Commonwealth of Learning & Athabasca University (2014)
8. Sharples, M., Taylor, J., Vavoula, G.: A theory of learning for the mobile age (2005). Disponible en: http://www.lsri.nottingham.ac.uk/msh/ Papers/Theory%20of%20Mobile%20Learning.pdf
9. Parsons, D., Ryu, H., Cranshaw, M.: A desing requirements framework for mobile learning environments. Journal of computers, 2(4), 1-8 (2007)
10. Koole, M.: A model for framing mobile learning. Athabasca University, Canada (2009). Disponible en: http://auspace.athabascau.ca:8080/dspace/bitstream/2149/2016/1/02_Mohamed_Ally_2009-Article2.pdf
11. Kearney, M., Schuck, S., Burden, K., Aubusson, P.: Viewing mobile learning from a pedagogical perspective. The Journal of the Association for Learning Technology (ALT), 20 (2012)
12. Aguilar, G., Chirino, V., Neri, L., Noguez, J., Robledo-Rella, V.: Impacto de los recursos móviles en el aprendizaje. 9ª Conferencia Iberoamericana en Sistemas, Cibernética e Informática, Orlando Florida, EE.UU (2010)
13. Wang, M., Shen, R., Novak, D., Pan, X.: The impact of mobile learning on students' learning behaviours and performance: Report from a large blended classroom. British Journal of Educational Technology, 40(4), 673-695 (2009)
14. Ramos, A. I., Herrera, J. A., Ramírez, M. S.: Desarrollo de habilidades cognitivas con aprendizaje móvil: un estudio de casos. Redalyc, XVII(34), 201-209 (2010)

15. UNESCO: Turning on Mobile Learning in Latin America. Illustrative Initiatives and Policy Implications. Working Papers Series on Mobile Learning (2012)
16. Cuesta, M., Herrero, F.: Introducción al muestreo. Depto. de Psicología, Universidad de Oviedo (2010). Disponible en: http://www.psico.uniovi.es/DptoPsico logia/metodos/tutor.7/
17. Organista-Sandoval, J., Serrano-Santoyo, A.: Appropriation and Educational Uses of Mobile Phones by Students and Teachers at a Public University in Mexico. Creative Education, 5, 1053-1063 (2014)
18. Trinder, J.: Mobile technologies and systems. En A. Kukulska-Hulmey & J. Traxler (Ed.), Mobile Learfning: a handbook for educators and trainers pp. 7-24. London and New York: Routledge Taylor & Francis Group (2005)
19. Bustos, O.: Mujeres y Educación Superior en México: Recomposición de la matrícula universitaria a favor de las mujeres. Repercusiones educativas, económicas y sociales (2003)
20. Aguilar, J., Ramírez, N.: Hábitos de consumo de las tecnologías información en los estudiantes universitarios de Tijuana. Munich Personal RePEc Archive, 4718,07 (2006)
21. Castellano, N.: Los usos de los dispositivos móviles (*smartphones*) en la conformación de un entorno de aprendizaje personal: el caso de la Facultad de Ingeniería de UABC Ensenada. XXVI simposio internacional de computación en la educación. Monterrey, Nuevo León (2010)
22. Crovi, D., Garay, L. M., López, R., Portillo, M.: Uso y apropiación de la telefonía móvil. Opiniones de jóvenes universitarios de la UNAM, la UACM y la UPN. Revista científica de la Asociación Mexicana de Derecho a la Información, 3 (2011)
23. Woodcock, B., Middleton, A., Nortcliffe, A.: Considering the Smartphone Learner: an investigation into student interest in the use of personal technology to enhance their learning. Student Engagement and Experience Journal,1-1 (2012)

A Literature Review on the Use of Soft Computing in Support of Human Resource Management

Jorge A. Rosas-Daniel, Oscar M. Rodríguez-Elias,
Maria de J. Velazquez-Mendoza, and Cesar E. Rose-Gómez

División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Hermosillo, Ave. Tecnológico S/N, Hermosillo, México.

roda.jorgearmando@gmail.com;
{omrodriguez,rvelazqu,crose}@ith.mx

Abstract. Organizational Human Resource Management (HRM) presents complex problems, some of them difficult to aboard by traditional problem solving strategies, because of the necessity to face subjective or imprecise issues, such as the evaluation of knowledge or abilities of a candidate for a specific job. Because of the above, researchers and practitioners have used Soft Computing, or Computational Intelligence, approaches to aboard some of the problems faced in HRM. This paper present the results of a systematic literature review conducted to know how Soft Computing has contributed to solve the problems that exist in HRM. The research identified is classified according to the problems faced in the field of HRM, and the Soft Computing techniques used to aboard such problems.

Keywords: Soft Computing, Computational Intelligence, Human Resource Management, Systematic Review.

1 Introduction

Knowledge has become a very important factor for many organizations; in fact it is the main input in the development of the products or services provided by many organizations. The last has conducted to a growing interest in the field of knowledge management [1]. Several authors have observed that for knowledge management to be successful, such initiatives must focus on the human resources [1, 2]. As Drucker observed [3], human resources, and particularly those in charge of knowledge intensive activities, could become the most valuable resource for XXI century organizations.

Human resources (HR) are an essential and strategic part of every organization [4, 5]. Individuals who are members of an organization make decisions and define the goals and the road the organization will follow, they control and configure the technological systems, create and share the knowledge needed for the organization's growth, identify the problems that interfere with the performance of the individuals, and hence the whole organization, among many other things. By its nature, HR Management

(HRM) faces complex problems difficult to address with traditional problem solving strategies, for instance because of the need for using qualitative and quantitative multi -criteria to evaluate individuals. If we include the need for managing the knowledge of HR this becomes even more complex because of the subjective and imprecise nature of people's knowledge. These characteristics of HRM have taken researchers and practitioners to explore the use of Soft Computing or Computational Intelligence techniques into the different activities of HRM.

To know the manner in which Soft Computing techniques have contributed to the solution of some of the problems that HRM face, in this paper we present the results of systematic literature review on the use of Soft Computing to address the problems present in the HRM field. In this paper we describe the main results of such research; particularly the main problems addressed by Soft Computing techniques, and the main techniques used. The remains of this paper are structured as follows: first, section 2 introduces the basic concepts related to the research. In section 3, the methodology we followed for conducting the research is described. Section 4 presents the results obtained once the retrieved documents were analyzed and classified. The discussion of the observed in those results is in section 5, followed by our conclusions in section 6.

2 Theoretical Background

The present work is based on the relationships of two main fields: Human Resource Management and Soft Computing. In this section, we shall develop on these two concepts in order to contextualize our research.

2.1 Human Resource Management

HRM is about the efficient management of the working people in order to make them productive and satisfied workers [5]. Without HR, computer systems, payment plans, mission declarations, procedures or programs won't be needed. It is required that the members of the organization be harnessed in an efficient way, but taking care of their wellness and comfort. To maximize the use of human resources, different factors, either qualitative or quantitative, must be taken into account. These factors relate to the complexity of human nature and of the processes of an organization.

The traditional approaches in the field of HRM require personnel capable of following orders. Personnel is instructed on what they should do and how, they must follow stablished procedures. However, when jobs become more complex, they also become more intensive in the use of knowledge [2]. This requires people that not only know what to do, but people with abilities, skills, and capacities to make decisions about how to do their jobs, when to do it, under which conditions, and knowing why they should do it [3]. This last conducts to the necessity of not only follow the record of the academic information and the career path of individuals; it is required to go farther and get records of the skills, competencies, formal and informal knowledge, experience, values, culture, etc., in order to make a better adjustment of the candidates

to a job [6]. Hence, choosing the better HR for a vacancy in an organization, as well as maintaining them inside the organization could become a very complex task [7].

HRM include several activities such as staff planning, recruitment, selection, rewards management, job evaluation, performance management, capacitation planning, promotion, hygiene and security, among others [5]. These activities should be aligned to the factors that influence the organization's processes and personnel satisfaction to function correctly, in order to achieve the productivity and competitive success that the organizations pursue. The complexity of the individuals and processes of organizations, lead them to require the support of intelligent technologies to carry out their tasks in a better way, while making better use of their human resources. There is why Artificial Intelligence techniques have proven to be useful in support of HRM [8].

2.2 Soft Computing

Soft Computing, or Computational Intelligence, represents a set of techniques for information processing, useful in cases where traditional algorithmic techniques could not exists, or be too complex [9]. These techniques process and interpret data (numeric, symbolic, binary, logic, images, etc.) in connection with a symbolic representation of knowledge, by imitating some of the aspects of intelligent human behavior.

Lotfi Zadeh defines Soft Computing as "a collection of methodologies that aim to exploit the tolerance for imprecisions and uncertainty to achieve tractability, robustness, and low solution cost" [10]. Soft Computing include the use of techniques such as artificial neural networks, fuzzy logic, genetic algorithms, rough sets, probabilistic methods, among others. In summary, Soft Computing provides intelligent mechanisms that may be useful to solve some of the problems faced by HRM activities.

3 Methodology

The present study consisted on the realization of a systematic literature review following the procedure proposed in [11], where a systematic review is described as a mean to evaluate and interpret the relevant research related to a research question or phenomenon of interest, in order to get a reasonable evaluation by using a reliable, rigorous and auditable methodology [11].

The goals of the review were: to identify the problems associated to HRM in which Soft Computing has been used to provide a solution; to identify the Soft Computing techniques used to solve the problems faced in HRM; to identify how those techniques are being used to solve the problems of HRM.

Considering the above objectives, the review was carried out taking into account the following research questions: 1) In which HRM areas have been used Soft Computing techniques? 2) What HRM problems have been faced with Soft Computing techniques? 3) Which Soft Computing techniques have been used in support of HRM? 4) How those techniques have been used?

3.1 Planning the Review

The literature review was conducted by searching in digital databases. We developed a list of terms related to HRM and Soft Computing. The initial selection of them was made with Brainstorming. This initial list of terms was used for a preliminary search where retrieved documents were less than expected. Because of these, we contacted with an expert in the field of HRM to improve the list of terms, this expert helped us to extend the list of terms, and provided us some relevant literature to improve the list. The same was made in the field of Soft Computing. The terms for the HRM field were 67, and 7 for the field of Soft Computing. Additionally we defined a list of 5 synonyms for the word "employee", these synonyms were used to create new terms to get a wider search, for instance, "Employee Selection" and "Staff Selection".

The digital databases used for the review were selected based on the fact that they include journals in the field of Soft Computing or related to it, and because access to them was available to us. The selected databases were: Thomson Reuters Web of Science, EBSCO, IEEE Xplore, and Elsevier sciencedirect.

To conduct the review, first we opted for creating one general query. This was because we wanted to avoid debugging duplicated documents in a same database, as well as reducing search time. To create this query, we grouped the terms into two groups: the first one was a set composed of the union of terms in the Soft Computing field, and the second one the set resulting of the union of the terms in the HRM field. The final query was the intersection of the two sets. The documents were filtered using three aspects of them: the title, the abstract, and the keywords.

While using this general query, we faced one big problem: the search engine of each database present technical differences that made it impossible for us to use the general query on all of them. Some search engines presented restrictions on the number of terms that can be used in a query therefore they rejected it. To solve this, it was necessary to remake the search method by dividing the general query into several queries; each query was made by the intersection of each term of the HRM field as the first part of the query, and the set of terms in the field of Soft Computing as the second one. Thus, there were 67 queries on each database, 268 totally. Additionally, it was necessary to use an additional filter in the IEEE database to reduce the documents retrieved by searching just in the range of time of 2004 to 2014. This was because the search engine was not filtering the documents correctly according to the query, and therefore it retrieved too much documents, most of them irrelevant.

3.2 Inclusion and Exclusion Criteria

To establish the criteria to be used to determine the documents to be considered is an important part since it gives validity to results of the review [11]. The inclusion and exclusion criteria that were considered are the following:

- The language of the document must be English or Spanish.
- The title must include a reference to the HRM and Soft Computing fields.
- If the title does not include explicitly a reference to the HRM field, but it can be inferred that it is the application area of a Soft Computing technique, or it is not to-

tally clear, we will search for an explicit reference of the HRM field in the abstract or keywords.
- If the title does not include an explicit reference to the Soft Computing field, but it does with the HRM field, we will search for an explicit reference of the Soft Computting field in the abstract or keywords.
- If the title does not include an explicit reference of any of both fields, we will search for an explicit reference of both fields in the abstract and keywords.
- If none of the above criteria are fulfilled, the document is rejected.

4 Results

Table 1 shows the number of documents that were found and selected. The documents are organized using the HRM area where the Soft Computing techniques were applied as the rows, and the Soft Computing technique used as the columns. The column labeled others, groups the documents using combinations of the four techniques, as well as one document using Rough sets, and another that uses an imperialist competition algorithm. The row named others, groups the documents in which Soft Computing techniques were used in the areas of HRM with less than four incidences (most of them with just one or two).

Table 1. Results of the review (GA = Genetic Algorithms, FL = Fuzzy Logic, ANN = Artificial Neural Networks, NFN = Neuro-Fuzzy Networks).

Problem	GA	FL	ANN	NFN	Others	Total
Personnel Prediction			7	2	1	10
HRM Evaluation		1	3			4
Job Scheduling	8	2	1		3	14
Personnel Assessment	1	10	5	2		18
Personnel Assignment or Turnover	17				3	20
Personnel Selection		14	3	2	1	20
Others		5	5	1	4	15
Total	26	32	24	7	12	101

In the following subsections we briefly describe the problems of HRM faced with Soft Computing, and the techniques used to solve such problems.

4.1 HRM Problems Faced with Soft Computing

Personnel Selection. It can be observed on table 1 that personnel selection is one of the two most faced areas in the field of HRM. The problems faced are related to the selection of the best individual for the job. This problem uses to be addressed from a

multicriteria point of view, where the individuals are evaluated with respect to several essential capacities, abilities or attributes for a good performance into the specified job [12–14]. In these cases, fuzzy logic is used for addressing the uncertainty related to the description of the profile of a specific job, or related to the evaluation of the profile of a specific individual. Some times the selection is focused on determining the best of all the candidates [15]; others in searching for an ideal candidate [16]; and in other cases on defining a solution that helps avoid certain undesired behaviors [17]. Fuzzy Logic has also been used combined with neural networks to construct learning systems capable of identifying the better characteristics to describe a specific job [18].

Personnel assignment and tournover is the other most faced problem. Genetic algorithms have been used mainly in this area, where it is required to find the better combination of individuals within an organization, in order to maximize or minimize one or several characteristics, such as cost, time, quality, etc. [19, 20]. Personnel reassignment can help to eliminate repetitive tasks by using ergo-nomic, competencies, environment and physical abilities [21], in the search of a better performance, and to avoid diseases. Another use has been in the problem of fair payment with the use of personnel assignment, and evaluation and assigning projects to employees to avoid dissatisfaction because of an unfair payment or economic differences [22]. Additionally, in [23] a methodology for addresing the decreased efficiency caused by the "learning-forgetting" efect is proposed.

Personnel assignment can be also addressed as the formation of work teams. Selecting the better team for a job is a difficult task. One must take into account the abilities and the relationships of the candidates to fulfill a specific task: "good teams before good individuals" [24].

Personnel assessment looks for defining the value that individuals have for the organization [25–27]. For addressing this problem, it has been proposed the use of neural networks for talent discovery [28]. Fuzzy logic has been used to evaluate individual performance considering multiple sources [29], and to consider employee satisfaction, learning and performance in capacitation events [30]. Neural networks were also used to predict and optimize personnel efficiency by valuating its attributes in order to identify those that are better to determine the impact of the individuals into the organizational efficiency [31, 32]. Additionally, neuro-fuzzy networks have been used to evaluate the individuals that resigned to a job in order to predict absenteeism [33].

Job schedulling consists on asigning resources to a job, such as HR, considering diferent restrictions, criteria, and one or several goals. In [34], HR are assigned acording to their skills, by searching for the best individual and the best moment to avoid delays in tasks and to reduce perturbations while inserting new preventive or corrective maintenance tasks. In [35] neural networks and genetic algorithms are combined to consider human factors in resources and activity scheduling, in this work, authors used past schedulings as a mean to obtain information to define new schedules using neural networks as the learning method.

There are also important works in other areas, such as the job classification according to risk of physical disorders by using neural networks [36], in order to take decisions that allow avoid such disorders. As well, in [37] authors used an imperialist competition algorithm to solve the problem of positioning HR in the adequate job to

obtain cost and time reductions. Neuro fuzzy approaches have also been used to discover implicit knowledge to allow predicting the future performance of employees, in order to select the best individual for a project or a position [38].

In summary, considering the Soft Computing techniques, the most used to face problems in the field of HRM seems to be fuzzy logic, which has been mostly used to address problems related to personnel selection and assessment. On the other hand, genetic algorithms have been mainly used for solving problems related to personnel assignment and rotation, and for job scheduling with different focus. Finally, Artificial Neural Networks have been the most widely used technique if we consider the number of different areas of the HRM where they have been applied.

From the results of the review we can observe that it is clear that Soft Computing has had a significant impact in the HRM field. It has proven to be a useful tool for solving complex problems presented in almost all the activities of the HRM process. Nevertheless, even though we conducted a wide review, we consider that the works found in literature are still a few, and more research should be carried out in order to exploit the benefits that Soft Computing could bring to solve many of the problems faced in HRM, an essential process for every organization.

5 Discussion

In the literature we found that some of the main problems addressed with soft computing techniques in HRM are personnel assessment and selection, as well as the job assignment of personnel, using muticriteria approaches. In these problems it is common to search for the individual that better fits to an ideal to fulfill a specific job, by giving it a qualification from the best to the worst, or from 0 to 1. Fuzzy logic is the main mechanism used to assign those qualifications, since it is commonly needed the use of several attributes that are frequently subjective and imprecise.

However, we observed that the methods used do not allow to know if individuals are overqualified, an important issue to make the best decision while assigning personnel to a job. We also observed that most works reviewed use to consider all the criteria as if their where just one, since they sum them before making the match between the individual and the job, hiding the qualification of specific criteria.

Multicriteria approach is robust, however, it could be possible to obtain important information to make a better decision if we would be able to show the result using a multilevel approach, for instance, by creating a job profile based on the activities tha should be performed, and describe these activities according to the competencies required to accomplish them. In this way, we could evaluate individuals according to their competencies, but at the same time know what activities or jobs can do in a satisfactory manner, and which ones do not.

In [39, 40] an approach to describe job and employees profiles is proposed. It allows realizing the description in a multilevel way, describing the abilities, knowledge, competencies, and values required for a job or role of an organization. This proposal has been defined following a knowledge representation approach based on traditional logic. Therefore, it does not allow facing the uncertainty that exists in a natural form

in the description of the necessities of a job, as well as on the description of the profile of a candidate for such a job. The systematic review conducted has helped us to identify the manner in which fuzzy logic can help us to aboard these limitations.

As a result of the above, and as a continuation of the present work, we will search for a more detailed review of the manner in which fuzzy logic has been used to solve the problems present in the activities of personnel selection, assessment and assignment, in order to adapt the proposal of [39, 40].

We expect that such adaptation will facilitate the use of a description of job and employee profiles to help the analysis and the matching of them in order to obtain a comprehensive solution for the several problems faced by HRM, including:

- Facilitate Job analysis.
- Select the best candidates for a job.
- Improve the assignment of personnel to a job.
- Improve the design of career plans for the employees of an organization.
- Improve capacitation programs and planning according to the specific needs of each job or employee.

Job analysis is an important activity for HRM. It allows obtaining information about the relevant issues of a job, giving as main result a description (summary) and specification (knowledge, abilities, competencies, skills, etc.) of it [5]. The above is essential to carry out in a better way the rest of the activities of HRM. There is why as part of our further work we will focus on providing a mechanism to facilitate job analysis through an approach based on fuzzy logic. As observed in the table 1, it is an area of research that has not been explored from this point of view in literature.

6 Conclusions

Soft Computing has offered to HRM different alternative solutions to the problems that managers face while they work with the complexity of human resources. The use of fuzzy logic to represent the uncertainty given by the human complexity, genetic algorithms used in the combinatory optimization problems present in job and tasks assignment, and neural networks to classify, evaluate, and predict conducts of individuals and organizations, within others, have demonstrated to be potent and efficient mechanisms to solve some of the most complex problems faced in HRM.

Most of the works found in literature focus on solving specific and particular problems. It is needed a more comprehensive approach capable of impacting in most activities of HRM. There exist important problems in HRM that have been not addressed yet by the use of Soft Computing techniques. We consider that one alternative to aboard these deficiencies is to provide a solution for the job analysis phase that facilitate the use of Soft Computing techniques in the rest of the HRM activities. Taking these into account, as a result of the present work we have proposed to adapt an approach to describe jobs and employees profiles, in order to such profiles to be described following a schema that facilitate the use of fuzzy logic in the analysis of

jobs. To accomplish this, we will make a more detailed analysis of the manner in which fuzzy logic has been used in the field of HRM.

Finally, it is important to mention that even though we consider that the result of the review is a good approach to what is reported in the literature about the use of Soft Computing in HRM, a more extended review could be made in order to consider other important sources, such as the digital database of Springer, ACM, and others that were not considered in our review.

Acknowledgment. This research was partially supported by PROMEP (grant PROMEP/103.5/12/46333), and by DGEST (grant 513.1/2171/2014).

References

1. Dalkir, K.: Knowledge Management in Theory and Practice. The MIT Press (2011).
2. Wiig, K.: People-focused knowledge management: how effective decision making leads to corporate soccess. Elsevier, Amsterdam (2004).
3. Drucker, P.F.: Management Challenges for the 21st Century. HarperBusiness (2001).
4. Armstrong, M.: Strategic Human Resource Management: a guide to action. Kogan Page, London (2006).
5. Ivancevich, J.M.: Administración de Recursos Humanos. McGraw Hill, México, D.F. (2005).
6. Breaugh, J. a.: Research on Employee Recruitment: So Many Studies, So Many Remaining Questions. J. Manage. 26, 405–434 (2000).
7. Abraham, J., Morin, L., Renaud, S., Saulquin, J., Soparnot, R.: What do Experts Expect from Human Resource Practices. Glob. J. Bus. Res. 7, 121–133 (2013).
8. Martinsons, M.G.: Knowledge-based systems leverage human resource management expertise. Int. J. Manpow. 16, 17–35 (1995).
9. Rutkowski, L.: Computational Intelligence. Springer Berlin Heidelberg, Berlin, Heidelberg (2008).
10. Zadeh, L.A.: Soft computing and fuzzy logic. IEEE Softw. 11, 48–56 (1994).
11. Kitchenham, B.: Procedures for Performing Systematic Reviews. , Eversleigh, Australia (2004).
12. Güngör, Z., Serhadlıoğlu, G., Kesen, S.E.: A fuzzy AHP approach to personnel selection problem. Appl. Soft Comput. 9, 641–646 (2009).
13. Canós, L., Liern, V.: Soft computing-based aggregation methods for human resource management. Eur. J. Oper. Res. 189, 669–681 (2008).
14. Zhang, S., Liu, S.: A GRA-based intuitionistic fuzzy multi-criteria group decision making method for personnel selection. Expert Syst. Appl. 38, 11401–11405 (2011).
15. Keršulienė, V., Turskis, Z.: Integrated fuzzy multiple criteria decision making model for architect selection. Technol. Econ. Dev. Econ. 17, 645–666 (2011).
16. Kelemenis, A., Ergazakis, K., Askounis, D.: Support managers' selection using an extension of fuzzy TOPSIS. Expert Syst. Appl. 38, 2774–2782 (2011).
17. Kabak, M., Burmaoğlu, S., Kazançoğlu, Y.: A fuzzy hybrid MCDM approach for professional selection. Expert Syst. Appl. 39, 3516–3525 (2012).
18. Shahhosseini, V., Sebt, M.H.: Competency-based selection and assignment of human resources to construction projects. Sci. Iran. 18, 163–180 (2011).

19. Mungle, S., Benyoucef, L., Son, Y.-J., Tiwari, M.K.: A fuzzy clustering-based genetic algorithm approach for time–cost–quality trade-off problems: A case study of highway construction project. Eng. Appl. Artif. Intell. 26, 1953–1966 (2013).
20. Vanucci, S.C., Carrano, E.G., Bicalho, R., Takahashi, R.H.C.: A modified NSGA-II for the Multiobjective Multi-mode Resource-Constrained Project Scheduling Problem. 2012 IEEE Congress on Evolutionary Computation. pp. 1–7. IEEE (2012).
21. Mondal, P.K., Ahsan, A.M.M.N., Quayum, K.A.: An approach to develop an effective job rotation schedule by using genetic algorithm. 2013 International Conference on Electrical Information and Communication Technology (EICT). pp. 1–5. IEEE (2014).
22. Vasant, S.R., Vasant, A.R., Jani, N.N.: A Novel Proposed Approach for an Equitable Payment Distribution Using Genetic Algorithm. 2012 International Conference on Communication Systems and Network Technologies. pp. 848–851. IEEE (2012).
23. Yan, J., Wang, Z.: GA based algorithm for staff scheduling considering learning-forgetting effect. 2011 IEEE 18th International Conference on Industrial Engineering and Engineering Management. pp. 122–126. IEEE (2011).
24. Herrera, F., López, E., Mendaña, C., Rodríguez, M.A.: A linguistic decision model for personnel management solved with a linguistic biobjective genetic algorithm. Fuzzy Sets Syst. 118, 47–64 (2001).
25. Huang, M., Mou, R.: Improvement on the genetic algorithm and its application in employee performance evaluation. 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE). pp. V4–253–V4–256. IEEE (2010).
26. Moon, C., Lee, J., Lim, S.: A performance appraisal and promotion ranking system based on fuzzy logic: An implementation case in military organizations. Appl. Soft Comput. 10, 512–519 (2010).
27. Roland, B., Di Martinelly, C., Riane, F., Pochet, Y.: Scheduling an operating theatre under human resource constraints. Comput. Ind. Eng. 58, 212–220 (2010).
28. Waheed, S., Zaim, A.H., Zaim, H., Sertbas, A., Akyokus, S.: Application of neural networks in talent management. 2013 International Conference on Electrical Information and Communication Technology (EICT). pp. 1–4. IEEE (2014).
29. Hosseininezhad, F., Nadali, A., Balalpour, M.: A Fuzzy Expert System for performance evaluation of HRM with 360 degree feedback approach (Case study: An Iranian IT company). 2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT). pp. 487–491. IEEE (2011).
30. Esquivei García, R., Félix Benjamín, G., Bello Pérez, R.: Evaluación del impacto de la capacitación con lógica difusa. Ingeniare. Rev. Chil. Ing. 22, 41–52 (2014).
31. Azadeh, A., Saberi, M., Jiryaei, Z.: An intelligent decision support system for forecasting and optimization of complex personnel attributes in a large bank. Expert Syst. Appl. 39, 12358–12370 (2012).
32. Azadeh, A., Saberi, M., Moghaddam, R.T., Javanmardi, L.: An integrated Data Envelopment Analysis–Artificial Neural Network–Rough Set Algorithm for assessment of personnel efficiency. Expert Syst. Appl. 38, 1364–1373 (2011).
33. Martiniano, A., Ferreira, R.P., Sassi, R.J., Affonso, C.: Application of a neuro fuzzy network in prediction of absenteeism at work. 2012 7th Iberian Conference on Information Systems and Technologies (CISTI). pp. 1–4. IEEE, Madrid, Spain (2012).
34. Marmier, F., Varnier, C., Zerhouni, N.: Proactive, dynamic and multi-criteria scheduling of maintenance activities. Int. J. Prod. Res. 47, 2185–2201 (2009).
35. Zuters, J.: An Adaptable Computational Model for Scheduling Training Sessions. Annual Proceedings of Vidzeme University College "ICTE in Regional Development." pp. 110–113. , Valmiera, Latvia (2005).

36. Zurada, J., Karwowski, W., Marras, W.S.: A neural network-based system for classification of industrial jobs with respect to risk of low back disorders due to workplace design. Appl. Ergon. 28, 49–58 (1997).
37. Laleh, E., Lotfi, S., Isazadeh, A., Masodi, Y.: An Algorithm of Developed Imperialist Competition for Suitable Human Resource Layout. 2013 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). pp. 11–16. IEEE, Honolulu, HI (2013).
38. Huang, M.-J., Tsou, Y.-L., Lee, S.-C.: Integrating fuzzy data mining and fuzzy artificial neural networks for discovering implicit knowledge. Knowledge-Based Syst. 19, 396–403 (2006).
39. Velázquez Mendoza, M.J., Rodriguez-Elias, O.M., Rose Gómez, C.E., Meneses Mendoza, S.R.: Perfiles de Conocimiento en la Gestión del Recurso Humano de las Organizaciones. Congr. Int. Investig. Acad. JOURNALS. 4, 3209–3214 (2012).
40. Velázquez Mendoza, M.J., Rodriguez-Elias, O.M., Rose Gómez, C.E., Meneses Mendoza, S.R.: Modelo para diseño de perfiles de conocimiento: una aplicacion en la industria generadora de energía eléctrica. Res. Comput. Sci. 55, 125–135 (2012).

Reviewing Committee

Alfonso Alba Cadena
Ana Lilia Sosa López
Angel H. Corral Domínguez
Ariel Quezada Piña
Carelia Gaxiola Pacheco
Dora-Luz Flores
Everardo Gutiérrez López
Francisco E. Martínez Pérez
Gabriel Zepeda Martínez
Gabriela Noemí Aranda
Héctor Pérez Meana
Jorge Flores Troncoso
José Luis Briseño
José S. Murguía Ibarra
Julio Cesar Ponce
Mabel Vázquez Briseño
Margarita G. Mayoral Baldivia
Martin Olguín Espinoza
Norma Verónica Ramírez Pérez
Rafael Asorey Cacheda
Saúl Martínez
Verónica Quintero Rosas
Xiomara Zaldivar

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
noviembre de 2014
Printing 500 / Edición 500 ejemplares