

A Real Time and Robust Facial Expression Recognition and Imitation approach for Affective Human-Robot Interaction Using Gabor filtering

Felipe Cid, José Augusto Prado*, Pablo Bustos and Pedro Núñez

Abstract—Facial expressions are rich sources of communicative information about human behavior and emotion. The robot’s abilities to recognize and imitate emotions are powerful signals within Human Robot Interaction. This paper presents a real time system for recognizing and imitate facial expressions in the context of affective Human Robot Interaction. The proposed approach achieves a fast and robust facial feature extraction based on consecutively applying filters to the edge image. An efficient Gabor filter is used in this paper, joint to a set of morphological and convolutional filters to reduce the noisy and the light dependence of the image acquired by the robot. Then, a set of invariant edge-based features are extracted and used as input in a Dynamic Bayesian Network. The proposed system estimates the human emotion by recognizing different facial expressions using this Bayesian approach. The output of this classifier updates a geometric robotic head model, which is used as a bridge between the human expressiveness and the final robotic head. In this paper, the human facial expressions are successfully imitated by *Muecas*, a 12 degrees of freedom robotics head. Experimental results demonstrate the accuracy and robustness of the proposed approach compared to similar facial recognition systems.

I. INTRODUCTION

In the last decades, Human-Robot Interaction (HRI) has become an increasingly and growing research area at the intersection of other scientific fields, such as psychology, ethology and cognitive science. Most of the HRI methodologies are focused on non-invasive techniques based on the well-known natural language, just like people interact with each other in their daily life. Natural language, in conjunction with visual information, is a very efficient method for an interaction paradigm with robots (Fig 1). Body language, gestures or facial expressiveness have been successfully used for enhancing the empathy, the attention or understanding of social skills in a human-machine interaction [1], [2].

In a near future, a variety of different robots will inhabit human environments, such as homes or offices. Current robot’s abilities for an efficient communication with people in these real environments are very limited; nevertheless, these communication skills are the main key in the acceptance of robots by next human generations. Nowadays, design of social robots takes into account how to enhance the empathy and the attention [3], and also, how to include affective components in HRI. Some design alternatives, such

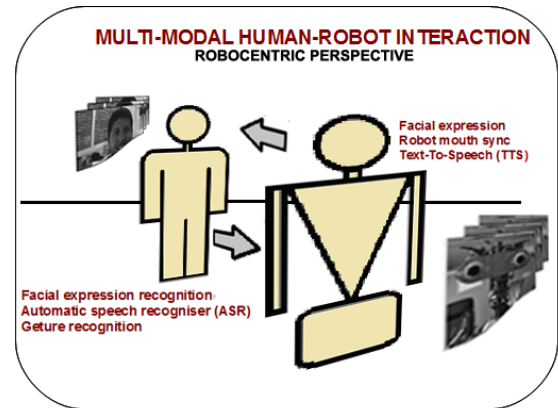


Fig. 1. Multi-modal HRI is usually based on visual and auditory information. Imitation is interpreted as revealing information to define and reinforce the relationship between human beings and robots.

as human shaped robots, decreases the gap between robot and human, during social interactions. Such user-friendly designs were extensive explored on the development of platforms for affective HRI (e.g., Kismet[4], Saya [5] or WE-4RII [6] robotics heads). In this respect, facial expression recognition and mimicry play an important role in human interaction and nonverbal communication.

In order to have an efficient affective interaction, not only the robot’s shape is determinant, but also the robot ability to acquire knowledge about the context or the emotional state of the interlocutor. Faces are rich and powerful sources of communicative information about human behavior and emotion. Thus, a large number of facial expression recognition systems are used in the literature as a first stage in affective HRI [7]. These systems provide support to the emotional responses of a robot inside a social dialogue through audio media or visual aids, creating a feedback for the content of the dialogue [8].

Human behavior imitation has been used for learning tasks and for enhancing the human-robot communication [9]. From a communication theoretical perspective, mimicry systems have been interpreted as revealing information to define and reinforce the relationship between individuals [11]. For the robot to express a full range of emotions and to establish a meaningful communication with a human being, facial expressions are vital.

The proposed approach presents a real time emotion recognition and imitation system based on facial expression analysis. On one hand, the facial expression recognition sys-

F. Cid, P. Bustos and P. Núñez are member of Robotics and Artificial Vision Lab. *Robolab* Group, University of Extremadura, Spain (e-mail: fecidb@alumnos.unex.es; pbustos, pnuntru@unex.es)

* J.A. Prado is with the Institute of Systems and Robotics, Dept. Electrical and Computer Engineering, University of Coimbra, Portugal (e-mail: jaugusto@isr.uc.pt)

tem consists of a robust feature extraction algorithm, which consecutively applies morphological and convolutional filters to reduce the noise and the dependence against changes of luminosity. After, an efficient Gabor filter is used for efficient edge detection. The output edge image of this filters bank is used for detecting and extracting scale-invariant facial features. Contrary to other approaches, these facial features are a combination of independent and antagonistic distortions of the face and constitute the input of a Dynamic Bayesian Network (DBN) used as classifier [10]. Four different emotions as well as non-emotional state are detected using this Bayesian approach (happiness, sadness, anger, fear and neutral). On the other hand, an imitation system is development and presented in this approach, where a robotic head model is used as a bridge to directly map from the detected emotion to the robot's actuators in safety conditions. This model is part of a cognitive module that is able to build selective representations of the self, the environment and the agents in it [12]. Finally, a set of experiments using a 12 degrees of freedom (DOF) robotic head has been achieved in order to present and comment the results of the recognition and imitation systems.

This paper is organized as follows: After discussing known approaches to facial expression and imitation techniques in Section II, Section III presents an overview of the proposed recognition and imitation system, which are described with details in Section IV and Section V, respectively. In Section VI, the experimental results are pointed out, and finally, Section VII describes the conclusions and future work of the presented approach.

II. RELATED WORKS

Automatic recognition of emotions has been studied in the last years for several authors. Current literature describes a complete system as multi-modal, that is, a recognition system that uses different information sources: face, gesture, body language, speech or physiological signals, amongst other [14], [13]. However, the majority of the approaches are based on the analysis of facial expressions using visual information. An interesting and updated review is shown in [15]. Common frameworks are based on the Facial Action Coding System (FACS) proposed by Ekman et al's [16]. In these works, the facial expressions are assigned to a small set of six prototypical expressions conveying the basic emotions.

Once a relevant face in the image is detected, two main problems arises in classical facial expression recognition systems: i) extracting the facial features; and ii) classifying the feature-based facial expressions into different categories. Detection and classification of facial features is a very diversified field in its classification ranges from the use of Active Appearance Models (AAM) [17], Support Vector Machines (SVM) [19] or Gabor filter bank [18]. Gabor filters have been commonly used for directly extracting features for recognition. However, it is computationally expensive. In [10] a method for detecting and classifying facial expression was proposed. It used color information and analysis of edges in order to extract facial features and use them as

input in a Dynamic Bayesian Network. Facial features in this work were very unstable, and dependent to the light condition. However, authors claim that the algorithm detects and recognizes four different facial expression in real time. The proposed approach is inspired on this work, being the results improved not only in robustness and accuracy, but also in processing time.

Several authors use robots in domestic environments with untrained users or people with disabilities [20], [21], [22]. In these works, authors achieve a natural HRI through the generation of facial expressions by the robot with the goal of maintaining a level of empathy and emotional attachment to robots [3]. These facial expression and emotion generation methods differs in the amount of facial expression that are possible to generate by the robot due to physical constraint [6]. In robotic heads with human-like characteristics such as the robotic head used in this paper, different works provide solutions for emotion generation depending on their physical constraints [23], [4]. This paper includes a model of the robotics head, as a module of a cognitive architecture that amongst other actions, prevents risk situations.

III. SYSTEM OVERVIEW

This paper presents a real time and robust facial expression recognition and imitation system. An overview of the proposed approach is illustrated in Fig. 2, which flows from left to right. Given an input video sequence S , the algorithm detects a relevant face in the robot's surrounding according to Viola and Jones' method [24]. Once the region of the face R has been estimate, the next step divides it into two different regions of interest, R_{top} and R_{bottom} . The proposed method incorporates a series of steps chosen to counter the effects of illumination variations, local shadowing and highlights. Next, morphological and convolutional filters, joint to a Gabor filter, have been also used to reduce noisy in the images. Facial feature extraction step detects and extracts invariant features from the edge image $f_i \in F$. These features and its time evolution are used as input in a Dynamic Bayesian Network, which classifies them into an emotion (happiness, sadness, anger, fear and neutral). Finally, after mapping the facial expression into the robotics face model, the it is imitated by the robotics head *Muecas*.

IV. FACIAL EXPRESSION RECOGNITION SYSTEM

This section describes with details the facial expression recognition system proposed in this paper. The algorithm uses the RGB sequence acquired by the robot, and process each image to detect a set of robust and invariant features in the user's face. These features are used as input of the Bayesian classifier proposed in [10]. Next, each stage illustrated in Fig. 2 is described.

A. Face detection

Let S being a sequence of RGB image acquired by the robot in a real interaction. A frame $I(t)$ of this sequence is then processed in order to detect a relevant face in the image. Viola and Jones' method is used in the proposed approach.

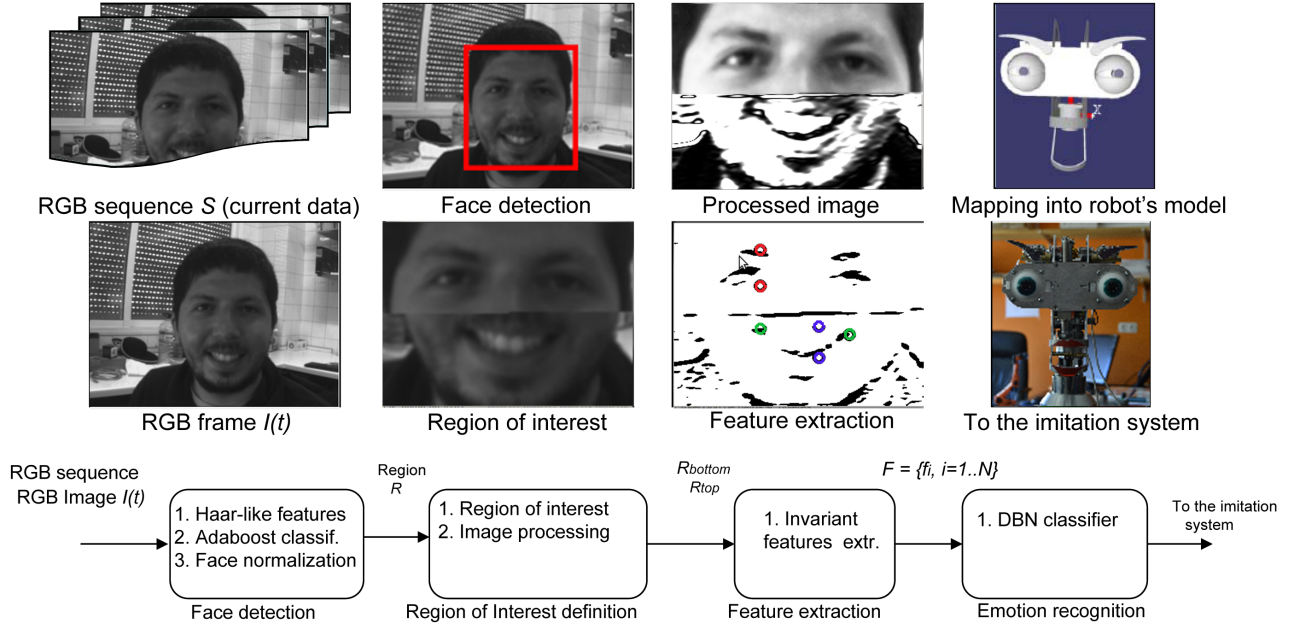


Fig. 2. Overview of the proposed methodology. The approach flows from left to right. See the text for more details.

This method uses Haar-like features and a cascade of simple classifiers [24] to detect the face in the image in real time. The relevant region R is then normalized to a fixed size and this image constitutes the input of the next stage.

B. Region of interest definition and pre-processing

Once the face has been detected, this image is processed in order to remove noisy and improve its light dependence. Besides, computational complexity is reduced by dividing the region R in two sub-regions of interest, R_{top} and R_{bottom} and converting the image in gray scale. These two regions are used to extract invariant facial features in the human face. In the proposed approach, nose is irrelevant for facial expression recognition, and then, it is removed from R . Thus, R_{top} and R_{bottom} are associated to the upper and lower region of the face, respectively. Let R being the face image of size $N \times M$, and let $p = (u, v)$ being the central pixel of this image, which is the approximated position of the nose in the image. Then R_{top} and R_{bottom} are defined as selective copies of R as follows: R_{top} of size $N \times (v - U_{Th})$ and R_{bottom} of size $N \times (v + U_{Th})$, where U_{Th} is an user-fixed threshold.

Next, the new image is converted to gray scale and processed. First, the proposed method incorporates a series of steps chosen to counter the effects of illumination variations while still preserving the main elements for use in recognition. The method is based on the approach described in [25]. The processing sequence follows a set of consecutive stages: i) gamma correction; ii) Difference of Gaussian (DoG) Filtering; iii) Masking; and iv) Contrast equalization.

Once the light dependence has been reduce, a set of morphological and convolutional filtering has been used. Noise in this image is associated to beard, wounds or similar elements, and it is mitigated applying Median, Blur and

Gaussian filters, consecutively. The final image is processed using a Gabor filter, which improves the edge-based facial feature detection.

C. Gabor Filtering

Gabor filter is a linear filter usually very effective and fast in the detection of edges with different orientations. This filter is used in the proposed approach as previous stage in the detection of facial features, which are extracted using the contours of the eyes, mouth or eyebrows in the face. Gabor impulse response in the spatial domain consists of a sinusoidal plane wave of some orientation and frequency, modulated by a two-dimensional Gaussian envelope. Let $I(u, v)$ being input image, then the output of the Gabor filter, $G(u, v)$, is given by:

$$G(u, v) = \exp\left(-\frac{1}{2}\left(\frac{u_\theta^2 + v_\theta^2}{\sigma^2}\right)\right) \cos\left(2\pi\frac{u_\theta}{\lambda} + \psi\right), \quad (1)$$

where θ , λ and ψ are associated to the sinusoidal plane wave (orientation, wavelength and phase, respectively), and being u_θ and v_θ described as:

$$\begin{aligned} u_\theta &= u \cos\theta + v \sin\theta \\ v_\theta &= -u \sin\theta + v \cos\theta \end{aligned} \quad (2)$$

Fig. 3(a) illustrates two different facial images. These images are processed according to the method described in this section. Results after applying light normalization and noise removal methods are shown in Fig. 3(b).

D. Invariant feature extraction

Feature extraction is a crucial step in facial expression recognition systems. The method described in this paper

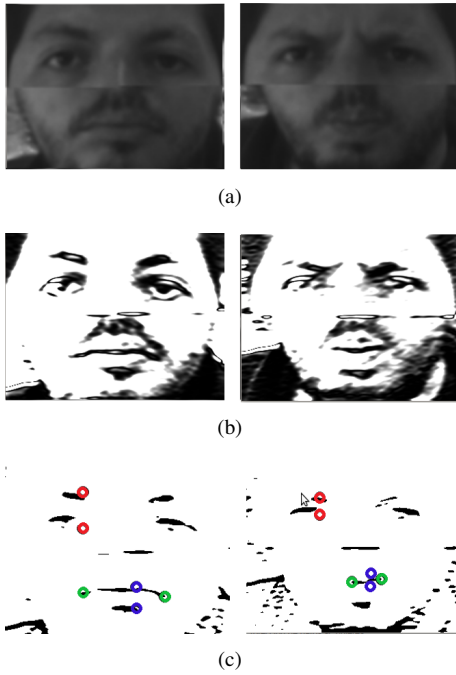


Fig. 3. Edge-based facial feature extraction: a) Region of interest in the face image (R_{top} and R_{bottom}); b) image after the pre-processing stage and Gabor filtering; and c) Features extracted in the image.

looks for a set of invariant edge-based features in the image $F=\{f_i \mid i = 1..N\}$, that is, features that are independent to the scale of the image or the distance from the user to the robot. Units are the basis of the proposed feature extraction algorithm. Each of these Action Units is a distortion on the face induced by small muscular activity, as it was described by the Facial Action Coding System (FACS). Different to other approaches, in this paper only a set of independent and antagonistic AUs are used (e.g., AU12 and AU15 in Fig. 4 are related to distortions in the lip corners, and they are antagonistic and independent). Only three features are defined in the edge face image, labeled as d_{eb} (red), d_{lc} (green) and d_m (blue) in Fig. 3(c), associated to the Euclidean distances between the upper contour of the eyebrows and the lower edge of the eyes, lip corners and upper and lower contour of the mouth, respectively. These values are easily detected analyzing the output of the Gabor filter as is illustrated in Fig. 3(c). In order to become independent to different scales or distances to the user, these edge-based features are always normalized using the values extracted in a neutral state.

E. Dynamic Bayesian Network for Facial Expression Recognition

The DBN takes advantage of the existing antagonism in some AUs to reduce the size of the dynamic Bayesian network. Thus, instead of using the 11 AUs as leaves for our DBN (Dynamic Bayesian Network), 7 variables are proposed as combinations of d_{eb} , d_{lc} and d_m . These variables are group the related antagonist and exclusive Action Units. The two-level network structure is illustrated in Fig. 5. The time influence that characterizes this network as a dynamic Bayesian network is also represented in Fig. 5.

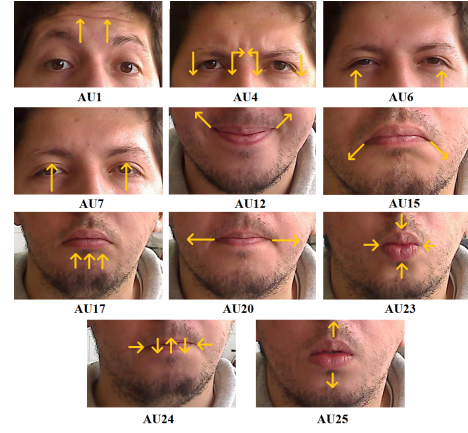


Fig. 4. Action Units (AUs) used in this paper.

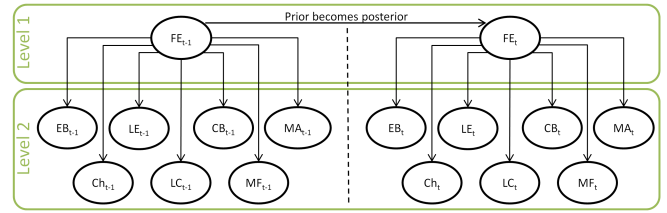


Fig. 5. Facial Expression Dynamic Bayesian network, two time intervals are shown.

In order to classify the Facial Expression (FE) produced by the user, a *dynamic Bayesian network's* is proposed, where the overall classification result achieved is the one foreseen by the belief variable FE , in the scope ($FE_{[neutral]}$, $FE_{[happy]}$, $FE_{[sad]}$, $FE_{[fear]}$, $FE_{[anger]}$).

Bayesian networks need to be supplied with learning data, the most common approach is to later use a threshold to find the matches when computing the probability of a new sample. In the proposed approach, to avoid the extant gaps, a pre-processing stage is achieved from the learning, setting up a Gaussian distribution that better fit into the data. The learning data was collected via supervised learning; seven random variables (leaf variables of our DBN) were gathered and manually related to the correct classification result.

The leaf random variables of our model, and it's respective virtual-scopes, are:

- EB : $\{AU1, AU4, none\}$; stands for *Eye-Brows*
- Ch : $\{AU6, none\}$; stands for *Cheeks*
- LE : $\{AU7, none\}$; stands for *Lower Eyelids*
- LC : $\{AU12, AU15, none\}$ stands for *Lips Corners*
- CB : $\{AU17, none\}$ stands for *Chin Boss*
- MF : $\{AU20, AU23, none\}$ stands for *Mouth's Form*
- MA : $\{AU24, AU25, none\}$ stands for *Mouth's Aperture*

In this approach, these 7 leaf variables are assumed to be independent given the facial expression (FE). Although some muscular movements from one area of the face may slightly affect other areas, this small influence could not be

detected by the cameras of the robot.

Thus, after the feature extraction process, the data (D) is obtained according to the following setup:

$$D = ((x_1, y_1) \dots (x_n, y_n)), x_i \in \mathbb{R}^d, y_i \in \mathbb{R} \quad (3)$$

Consider that y_1 to y_5 are the five possible emotional states ($FE_{[neutral]}$, $FE_{[happy]}$, $FE_{[sad]}$, $FE_{[fear]}$, $FE_{[anger]}$); and each dimension of x , corresponds to one of the previously described random variables, namely: EB , Ch , LE , LC , CB , MF and MA . Since the learning data may have gaps between its samples, a model is built assuming that (X_1, \dots, X_n) are independent given FE , and

$$X_i \sim N(\text{prior}^T x_i, \sigma^2) \quad (4)$$

At first, $\text{prior} \sim U(1/n)$, however throughout the iterations, the posterior of $t - 1$ becomes the prior on t .

Finally, by using the Bayes's rule, we have the posterior equation:

$$P(FE|x_m) = \frac{\prod_1^n P(x_i|FE) * P(FE)}{P(x_m)}, \quad (5)$$

where x_m is the most recent sensory data acquired.

The last dividend can be computed by using the Bayesian marginalization rule:

$$P(x_m) = \sum_{FE} \prod_1^k P(x_i|FE) * P(FE), \quad (6)$$

being $k = 7$, the number of random variables of the system.

This model include a dynamic convergence property over time. The resultant histogram from the previous frame is passed as prior knowledge for the current frame. The maximum number of frames for convergence has been limited to 5. If the convergence reaches a 80% threshold before 5 frames, the classification is considered complete (Fig 6). If not it keeps converging up to the fifth frame. If the fifth frame is reached and no value is higher than the threshold, the classifier selects the highest probability value (usually referred to as the *Maximum a posteriori decision* in Bayesian theory) as the classification result. The threshold is used as a control measure for the classification errors generated in the detection of the Action Units (AUs).

V. FACIAL EXPRESSION IMITATION SYSTEM

In this section, the proposed facial expression imitation system is described. The outline of the proposal is illustrated in Fig 7. Once the DBN recognizes the facial emotion, these are imitated by Iadex robotic Head *Muecas* (Fig. 8(c)). *Muecas* consists of 12 degree of freedom and it has been designed according to the human anatomy for generating facial expressions¹. Before mapping the facial emotion to the set of robot's actuators, the proposed imitation system performs internal simulations over the mesh model of the robotics head (Fig. 8(b)). Thus, assuming that all the characteristics of

¹For more information, you can visit www.robotlab.unex.es

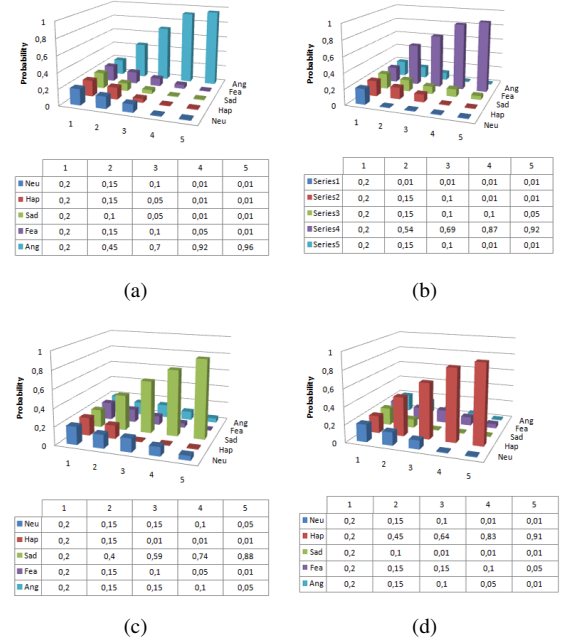


Fig. 6. Results from the classifier for different facial expressions.

Emotion	AUs	Muecas' component
Neutral	-	-
Happy	AU6-AU12-AU25	Eyebrows-Eyes-Mouth
Sad	AU1-AU4-AU15-AU17	Eyebrows-Eyes
Fear	AU1-AU4-AU20-AU25	Eyebrows-Mouth
Anger	AU4-AU7-AU17-AU23-AU24	Eyebrows

TABLE I

MAPPING OF FACIAL EXPRESSIONS (AUs) TO THE ROBOT'S MECHANICAL COMPONENTS.

each mechanical element are modeled, the system looks for possible collisions between them. In this respect, the system generates all the cinematic chain associated to the facial expression and if there is some collision, a retargeting of each mobile component is done to regenerate the cinematic chain.

Table I describes the relation between the AUs and the mobile elements of the mesh model of the robotic head *Muecas* (i.e., mapping) associated to each facial expression. In Fig. 8 different facial expression imitations are shown. Fig. 8(a) draws the facial expression recognized by the system (fear, anger, sad, happy and neutral, respectively). These facial expressions are previously mimicked by the robotic head model (8(b)) and by *Muecas*, after determining a collision-free configuration (Fig. 8(c)).

VI. EXPERIMENTAL RESULTS

In this section, a set of tests has been achieved in order to evaluate the performance of the proposed facial expression recognition and imitation system. Robustness and effectiveness are evaluated. Besides, a comparative study with the method proposed in [12] has been done. Here, edge-based features using color analysis was proposed. The proposed method has been evaluated using real video sequence. The

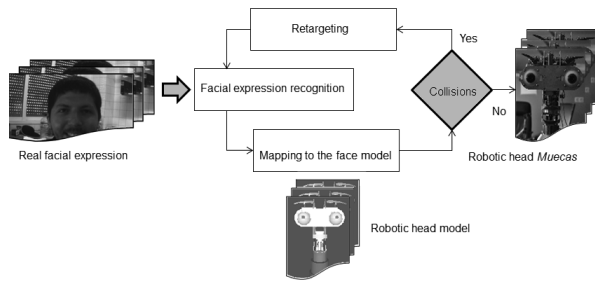


Fig. 7. Description of the Imitation system proposed in this paper.

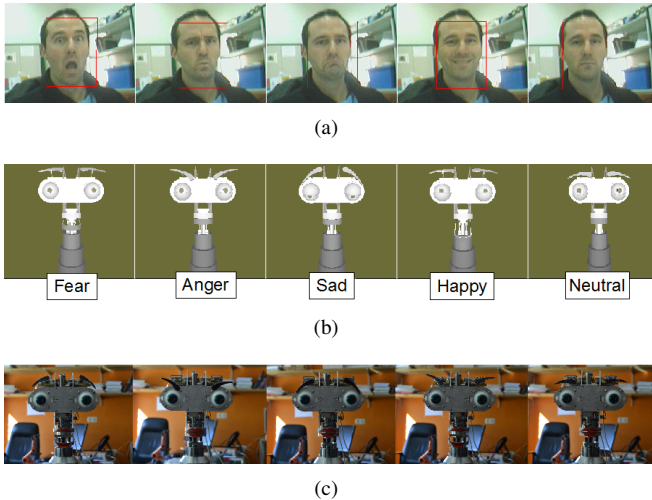


Fig. 8. a) Facial expression estimated by the recognition system; b) Mapping of the facial expressions into *Muecas* robotics head; and c) Facial expressions imitated on the real robotics head.

algorithms were developed in C++ and the benchmark tests were performed on a computer with a 2.8 GHz Intel(R) Core(TM) i7 CPU and 4Gb RAM running using GNU/Linux Ubuntu 10.10. Real data has been acquired using a firewire camera at 25 fps. The software to control the system is built on top of the robotics framework RoboComp [26]. Making use of the components and tools it provides and its communication middleware, an easy to understand and efficient architecture has been developed.

The proposed system is running on-line, acquiring and estimating the facial expressions in real-time. Thus, the system updates the emotional state at 12 or 24 milliseconds (processing time of the recognition system), and imitates the facial expression using the robotics head *Muecas* in real-time. The experiments consist of two tests, which were performed with untrained users in uncontrolled environments for quantifying the effectiveness of the proposed system. For each test, the user performed a series of continuous and randomness facial expressions. Besides, to verify the robustness of the system, the tests were performed with different lighting conditions and users of different facial features.

The first test is based on the estimate of the robustness of the facial expression recognition system. Through, the recognition of 20 different interlocutors with different facial

Test	Percentage of correctly detected facial expressions (r)
Sad	91%
Happy	94%
Fear	96%
Anger	95%
Neutral	96%

TABLE II
ROBUSTNESS OF THE FACIAL EXPRESSION RECOGNITION SYSTEM

Test	Percentage of improvement in the detection of FE (p_e)
Sad	18.4%
Happy	4.2%
Fear	15.1%
Anger	11.1%
Neutral	32.4%

TABLE III
PERCENTAGE OF IMPROVEMENT OF THE FACIAL EXPRESSION RECOGNITION SYSTEM

expressions and random sequence of emotions. Ground truth (*i.e.*, real facial expression) was selected by an expert. In Table II, the percentages of correctly detected facial expressions (r) are shown.

The second test is a achieved in order to compare the paper proposed in this paper with the method described in [12]. Identical conditions were used for evaluation the results after applying both two methods. Thus, the same database composed of 18 different interlocutors was used. Results of this comparison are shown in Table III, where the percentages of improvement in the detection of each facial expression p are illustrated. These results demonstrate an improvement in each detected emotion, especially in the neutral state. Besides, the proposed system improves the previous method in several aspects, such as: better performance in the detection of emotion, an error correction among states, facial expression recognition of the neutral state between expressions, greater accuracy in the recognition of facial features and the use of a smaller amount data on the trainer.

VII. CONCLUSIONS

In this paper, a real-time and robust facial expression recognition and imitation system for robotics head is proposed. Facial expression recognition stage is based on the use of invariant features on the edge image, which are robustly extracted thanks to a set of consecutive filters. Gabor filtering is used for an effective edge detection in the face image. Next, a Dynamic Bayesian Network is used for classifying these invariant features into an emotional state. The output of the Bayesian classifier is imitated by the robotic head *Muecas*. The system described in this paper was tested in order to verify the robustness, accuracy and improvement respect to other approaches of the recognition of facial expressions using different users, environments and lighting conditions.

Future works will be focused on a multi-modal interaction, where auditory information (*e.g.*, speech) will be used in

order to estimate the emotion. Besides, the proposed method will integrate RGBD information acquired by a low-cost sensor to estimate the facial expressions more accurately.

ACKNOWLEDGMENT

This work has been partially supported by the MICINN Project TIN2012-38079-C03-01.

REFERENCES

- [1] A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobreprez, S. Woods, C. Zoll and L. Hall, "Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents", In *Third International Joint Conference on Autonomous Agents and Multiagents Systems*, Vol. 1, pp. 194-201, New York, USA, 2004.
- [2] M. Siegel, C. Breazeal, and M. I. Norton, "Persuasive Robotics: The influence of robot gender on human behavior". In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2563-2568, October 2009.
- [3] A. Tapus and M.j. Mataric, "Emulating Empathy in Socially Assistive Robotics", In *AAAI Spring Symposium on Multidisciplinary Collaboration for Socially Assistive Robotics*, Stanford, USA, March 2007.
- [4] Kismet, Available at: <http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html>.
- [5] T. Hashimoto, S. Hitramatsu, T. Tsuji, and H. Kobayashi, "Development of the Face Robot SAYA for Rich Facial Expressions". In *Proc. of 2006 SICE-ICASE International Joint Conference Korea*, pp. 5423-5428, October 2006.
- [6] M. Zecca, T. Chaminade, M. A. Umilta, K. Itoh, M. Saito, N. Endo, "Emotional Expression Humanoid Robot WE-4RII- Evaluation of the perception of facial emotional expressions by using fMRI", In *Robotics and Mechatronics Conference ROBOMECH2007*, Akita, Japan, pp. 2A1-010, 2007.
- [7] Z. Zeng, M. Pantic, G. I. Roisman and T. Huang, "A Survey of Affect Recognition Methods: Audio, Visual and Spontaneous Expressions", In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, pp. 39 - 58, 2008.
- [8] T. Chen, "Audio-Visual Integration in multimodal Communication". In *IEEE Proceedings*, May, 1998.
- [9] R. W. Picard, "Affective Computing". MIT Press, pp. 88-91, 2000.
- [10] J. Prado, C. Simplicio, N. F. Lori, J. Dias. "Visuo-auditory Multimodal Emotional Structure to Improve Human-Robot-Interaction", in *International Journal of Social Robotics* Vol. 4, Issue 1, pp 29-51, 2012.
- [11] J. B. Bavelas, A. Black, C. Lemery, and J. Mullett, "I show how you feel: motor mimicry as a communicative act." *Journal of Personality and Social Psychology*, vol. 50, pp. 322-329, 1986.
- [12] F. Cid, J.A. Prado, P. Bustos and P. Nunez. "Development of a Facial Expression Recognition and Imitation Method for Affective HRI", in *Workshop of Physical Agents*, pp. 119-124, 2012.
- [13] A. Jaimes and N. Sebe, "Multimodal Human Computer Interaction: A survey", In *IEEE International Workshop on Human Computer Interaction in conjunction with ICCV 2005*, Beijing, China, 2005.
- [14] N. Sebe, I. Cohen, T. Gevers, T. S. Huang, "Multimodal Approaches for Emotion Recognition: A Survey", In *Internet Imaging VI, SPIE'05*, USA, 2005.
- [15] V. Bettadapura, "Face Expression Recognition and Analysis: The State of the Art", Tech Report, College of Computing, Georgia Institute of Technology, April 2012.
- [16] P. Ekman, W.V. Friesen, J.C. Hager, "Facial Action Coding System FACS", the manual, 2002.
- [17] K.-E. Ko, K.-B. Sim, "Development of a Facial Emotion Recognition Method based on combining AAM with DBN", in *International Conference on Cyberworlds 2010*, pp. 87-91, 2010.
- [18] H.-B. Deng, L.-W. Jin, L.-X. Zhen and J.-C. Huang. "A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA", *International Journal of Information Technology*. Vol.11, no.11. pp. 86-96, 2005.
- [19] S. S. Ge, C. Wang, C.C. Hang, "Facial Expression Imitation in Human Robot Interaction", In *Proc. of the 17 IEEE International Symposium on Robot and Human Interactive Communication*, Germany, pp. 213-218, 2008.
- [20] T. Mukai, S. Hirano, H. Nakashima, Y. Kato, Y. Sakaida, S. Guo, and S. Hosoe, "Development of a Nursing-Care Assistant Robot RIBA That Can Lift a Human in Its Arms", In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taiwan, pp. 5996-6001, October 2010.
- [21] M.J. Matarić, J. Eriksson, D.J. Feil-Seifer and C.J. Winstein. "Socially assistive robotics for post-stroke rehabilitation". In *Journal of NeuroEngineering and Rehabilitation*. 4:5. 2007.
- [22] C. Jayawardena, I. H. Kuo, U. Unger, A. Igc, R. Wong, C. I. Watson, R. Q. Stafford, E. Broadbent, P. Tiwari, J. Warren, J. Sohn and B. A. MacDonald. "Deployment of a Service Robot to Help Older People". In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems Taiwan*, pp. 5990-5995, October 2010.
- [23] T. Fukuda, M.-J. Jung, M. Nakashima, F. Arai and Y. Hasegawa, "Facial Expressive Robotic Head System for Human-Robot Communication and Its Application in Home Environment", In *Proc. of the IEEE*, Vol. 92, No. 11, pp. 1851-1865, 2004.
- [24] P. Viola and M. J. Jones, "Robust real-time face detection", in *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [25] X. Tan and B. Triggs. "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions", *IEEE Transactions on Image Processing*, Vol. 19, pp. 1635-1650, 2007.
- [26] L. Manso, P. Bachiller, P. Bustos, P. Nez, R. Cintas and L. Calderita. "RoboComp: a Tool-based Robotics Framework", In *Proceedings, SIMPAR Second International Conference on Simulation, Modeling and Programming for Autonomous Robots*. pp 251-262. 2010.