Electronic Letters on Computer Vision and Image Analysis 7(4):1-12, 2008

Facial Emotional Classifier For Natural Interaction

Isabelle Hupont, Eva Cerezo, Sandra Baldassarri

Departamento de Informática e Ingeniería de Sistemas,
Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza (Spain)
{478953, ecerezo, sandra}@unizar.es
Received 29th November 2007, Revised 26th February 2008, Accepted 3rd June 2008

Abstract

The recognition of emotional information is a key step toward giving computers the ability to interact more naturally and intelligently with people. We present a simple and computationally feasible method to perform automatic emotional classification of facial expressions. We propose the use of a set of characteristic facial points (that are part of the MPEG4 feature points) to extract relevant emotional information (basically five distances, presence of wrinkles in the eyebrow and mouth shape). The method defines and detects the six basic emotions (plus the neutral one) in terms of this information and has been fine-tuned with a database of more than 1500 images. The system has been integrated in a 3D engine for managing virtual characters, allowing the exploration of new forms of natural interaction.

Key Words: Face and gesture recognition, emotional classifier, multimodal interfaces

1 Introduction: facial emotional classification

Facial expression is the most powerful, natural and direct way between humans to communicate emotions, valuations and intentions. As pointed out by Bruce [1], human face-to-face communication is an ideal model for designing a multimodal human-computer interface (HCI).

A system capable of extracting emotional information from user's facial expressions would be of great interest for developing new interfaces which follow the human face-to-face communication model in the most realistic way. In particular, the creation of virtual environments populated by 3D virtual characters capable of understanding users' expressions and reacting accordingly represents, nowadays, a challenging but affordable task.

Nevertheless, to develop a system that interprets facial expressions is difficult. Three kinds of problems have to be solved: face detection in a facial image or image sequence, facial expression data extraction and facial expression classification (e.g. into emotional categories). This paper focus on the third problem: classification. This implies the definition of the set of categories we want to deal with, and the implementation of the categorization mechanisms.

Facial expression analyzers make use different methods of classification. The most commonly used are: patterns, neuronal networks or rules [8]. If a pattern-based method is used [2,3,4,25], the face expression found is compared with the patterns defined for each expression category. The best matching decides the classification of the expression. Most of these methods first apply PCA (Principal Components Analysis) and

Correspondence to: <478953@unizar.es >

Recommended for acceptance by Francisco Perales and Bob Fisher
ELCVIA ISSN: 1577-5097
Published by Computer Vision Center / Universitat Autonoma de Barcelona, Barcelona, Spain

LDA (Linear Discriminant Analysis) algorithms to reduce dimensionality. In the systems based on neuronal networks [5,6], the face expression is classified according to a categorization process “learned” by the neuronal network during the training phase. In general, the input to this type of systems is a set of characteristics extracted from the face (points or distances between points). The rule-based methods [7] classify the face expression into basic categories of emotions, according to a set of face actions previously codified. In the last years, other approaches have been developed based on Gabor filters [17,18], Hidden Markov Models [18,19], fuzzy logic [20,21], AdaBoost [24] or Support Vector Machines [22,23].

In any case, the development of automatic facial classification systems presents several problems. Most of the studies on automated expression analysis perform an emotional classification. The emotional classification of Ekman [9] is the most followed one. It describes six universal basic emotions: joy, sadness, surprise, fear, disgust and anger. Nevertheless, the use of Ekman’s categories for developing automating facial expression emotional classification is difficult. First, his description of the six prototypic facial expressions of emotions is linguistic and, thus, ambiguous. There is no uniquely defined description either in terms of facial actions or in terms of some other universally defined facial codes. Second, classification of facial expressions into multiple emotion categories should be possible (e.g. raised eyebrows and smiling mouth is a blend of surprise and happiness). Another important issue to be considered is individualization. The system should be capable of analyzing any subject, male or female of any age and ethnicity and of any expressivity.

The facial emotional we present in this paper uses a simple rule-based classification method capable to classify the facial expression of any user into one or blends of the 6 universal basic emotions of Ekman. The structure of the paper is as follows: in Section 2 our method is described; Section 3 presents the results obtained with static images and studies the influence of the race in the emotional classification results; Section 4 describes how the method is used to achieve emotional tracking in video sequences, allowing its integration in a 3D natural interface engine; finally, conclusions and comments about future work are discussed in Section 5.

2 A simple method for the automatic analysis of face expressions

Our method is based on the work of Hammal et al [10]. They have implemented a facial classification method for static images. The originality of their work consists, on the one hand, in the supposition that all necessary information for the recognition of expressions is contained in the deformation of certain characteristics of the eyes, mouth and eyebrows and, on the other hand, in the use of the Belief Theory to make the classification. Nevertheless, their method has important restrictions. The most important restriction comes from the fact that it is only able to discern 3 of the 6 basic emotions (without including the neutral one). This is basically due to the little information they handle (only 5 distances). It would not be viable, from a probabilistic point of view, to work with more data, because the explosion of possible combinations would remarkably increase the computational cost of the algorithm.

2.1 General description of the method

Our method studies the variation of a certain number of face parameters (distances and angles between some feature points of the face) with respect to the neutral expression. The objective of our method is to assign a score to each emotion, according to the state acquired by each one of the parameters in the image. The emotion (or emotions in case of draw) chosen will be the one that obtains a greater score.

For example, let’s imagine that we study two face parameters (P_1 and P_2) and that each one of them can take three different states (C^+, C^- and S, following the nomenclature of Hammal). State C^+ means that the value of the parameters has increased with respect to the neutral one; state C^- that its value has diminished with respect to the neutral one; and the state S that its value has not varied with respect to the neutral one. First, we build a descriptive table of emotions, according to the state of the parameters, like the one of the Table 1. From this table, a set of logical tables can be built for each parameter (Table 2). That way, two vectors of emotions are defined, according to the state taken by each one of the parameters (C^+, C^- or S) in a specific frame. Once the tables are defined, the implementation of the identification algorithm is simple. When a parameter takes a specific state, it is enough to select the vector of emotions (formed by 1's and 0's) corresponding to this state. If we repeat the procedure for each parameter, we will obtain a matrix of as many rows as parameters we study and 7 columns, corresponding to the 7 emotions. The sum of 1's present in each column of the matrix gives the score obtained by each emotion.

	P1	**P2**
Joy	C-	S/C-
Surprise	C+	C+
Disgust	C-	C-
Anger	C+	C-
Sadness	C-	C+
Fear	S/C+	S/C+
Neutral	S	S

Table 1. Theoretical table of parameters' states for each emotion.

Compared to the method of Hammal, ours is computationally simple. The combinatory explosion and the number of calculations to make are reduced considerably, allowing us to work with more information (more parameters) of the face and to evaluate the seven universal emotions, and not only four of them, as Hammal does.

		E1 joy	**E2 surprise**	**E3 disgust**	**E4 anger**	**E5 sadness**	**E6 fear**	**E7 neutral**
P1	**C+**	**0**	**1**	**0**	**1**	**0**	**1**	**0**
	C-	1	0	1	0	1	0	0
	S	0	0	0	0	0	1	1
		E1 joy	**E2 surprise**	**E3 disgust**	**E4 anger**	**E5 sadness**	**E6 fear**	**E7 neutral**
P2	**C+**	0	1	0	0	1	1	0
	C-	**1**	**0**	**1**	**1**	**0**	**0**	**0**
	S	1	0	0	0	0	1	1

Table 2. Logical rules table for each parameter.

2.2 Feature selection

The first step of our method consists of extracting the 20 feature points of the face that will later allow us to analyze the evolution of the face parameters (distances and angles) that we wish to study. Figure 1 shows the correspondence of these points with the ones defined by the MPEG-4 standard. The extraction of the points is made by means of a real-time facial feature tracking program described in [11].

The characteristic points are used to calculate the five distances shown in Figure 2. These five distances can be translated in terms of MPEG-4 standard, putting them in relation to the feature points shown in Figure 1 and with some FAPs (Facial Animation Parameters) defined by the norm. All the distances are normalized with respect to the distance between the eyes (MPEG Facial Animation Parameter Unit -FAPU- called

"ESo"), which is a distance independent of the expression. This way, the values will be consistent, independently of the scale of the image, the distance to the camera, etc.

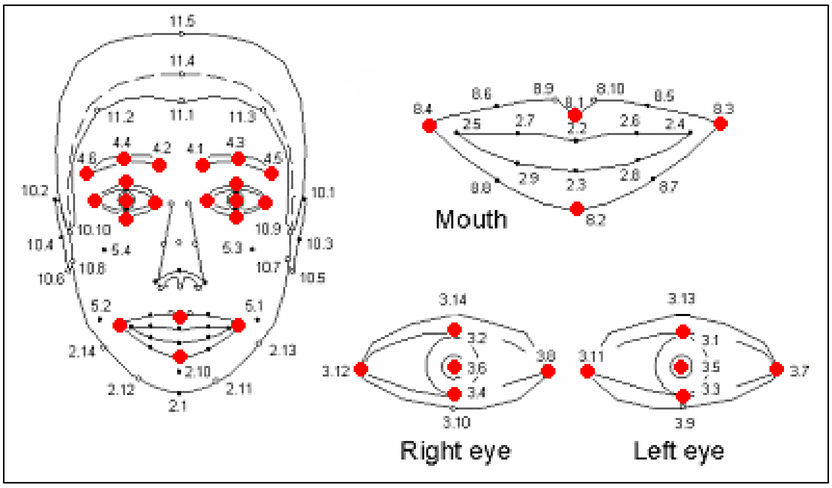
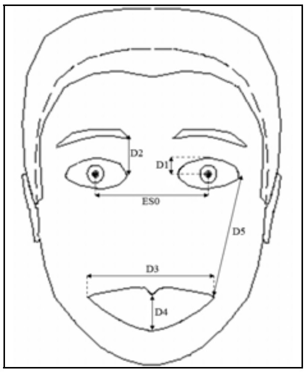


Fig. 1. Facial feature points used for the later definition of the parameters to analyze, according to MPEG-4 standard.



MPEG-4 FAPs NAME	FEATURE POINTS USED FOR DISTANCES
close_upper_l_eyelid close_lower_l_eyelid	D1=d(3.5, 3.1)
raise_r_i_eyebrow	D2=d(4.2, 3.8)
stretch_l_cornerlip stretch_r_cornerlip	D3=d(8.4, 8.3)
open_jaw	D4=d(8.1 , 8.2)
raise_r_cornerlip	D5=d(8.3, 3.7)

Fig. 2. Characteristic distances used in our method (left). On the right, relationship between the five characteristic distances and the MPEG-4 FAPs and feature points.

2.3 Database

In order to define the emotions in terms of the parameters states, as well as to find the thresholds that determine if parameter is in a state or another, it is necessary to work with a wide database. In this work we have used two different facial emotions databases: the FG-NET database [12] that provides video sequences of 19 different Caucasian people; and the MMI Facial Expression Database [13] that holds 1280 videos of 43 different subjects from different races (Caucasian, Asian and Arabic). Both databases show the 7 universal Ekmans' emotions.

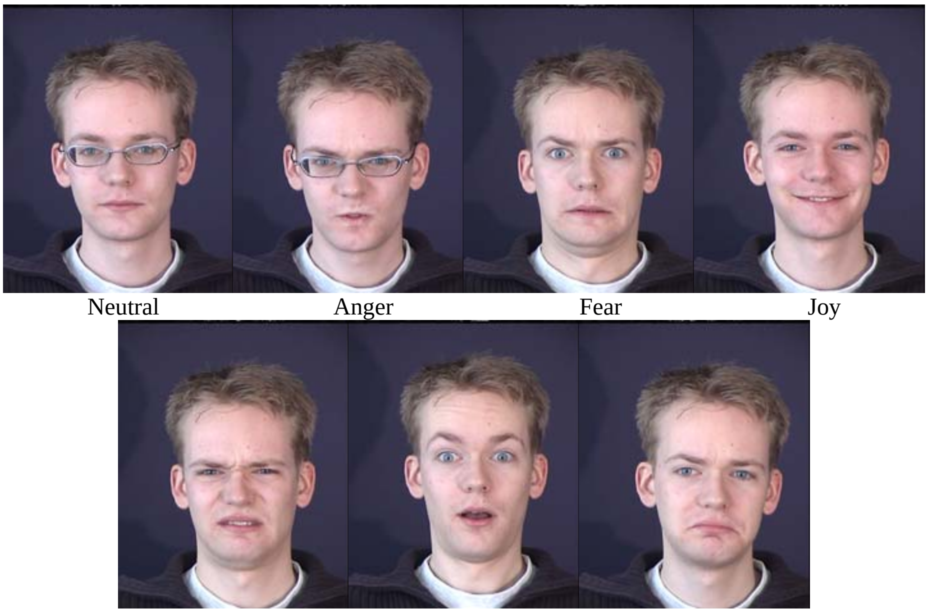


Fig. 3. Example of selected frames of the MMI Facial Expression Database [13].

3 Results

3.1 Initial results

First we considered to work with the same parameters as the Hammal's method, i.e. with the 5 characteristic distances shown in Figure 2. In order to build a descriptive table of each emotion in terms of states of distances, we must determine the value of the states of distances that define each emotion (C^+, C^- or S), as well as evaluate the thresholds that separate a state from another, for each distance. To do this, we studied the variation of each distance with respect to the neutral one, for each person of the database and for each emotion. An example of the results obtained for distance D_4 is shown in Figure 4. From these data, we can make a descriptive table of the emotions according to the value of the states (Table 3).

Note that the distances D1, D2 and D5 have a symmetric facial distance (one in each eye). Facial symmetry has been assumed after having calculated the high correlation between each distance and its symmetric.

The last step to complete our algorithm is to define the values of the thresholds that separate a state of another one, for each studied distance. Two types of thresholds exist: the upper threshold (marks the limit between neutral state S and state C^+) and the lower threshold (the one that marks the limit between neutral state S and state C^-). The thresholds' values are determined by means of automated cross-validation tests (90% of the database images are randomly chosen for training, 10% for benchmarking) over all the subjects and all the expressions of the databases we work with (a total of 1500 images aprox.). Those tests automatically generate possible combinations of thresholds depending on the statistics obtained for each distance (maximum, minimum, average, standard deviation and median values) and calculate the classification results for each combination. The finally chosen combination of thresholds is the one that maximizes the percentage of success in the emotional classification process. Figure 4 shows an example of thresholds estimation for the distance D_4.

In the evaluation of results, the recognition is marked as “good” if the decision is coherent with the one taken by a human being. To do this, we have made surveys to 30 different people to classify the expressions shown in the most ambiguous images. For example, in the image shown in Figure 5, the surveyed people recognized it as much "disgust" as "anger", although the FG-NET database classifies it like "disgust" exclusively. Our method obtains a draw.

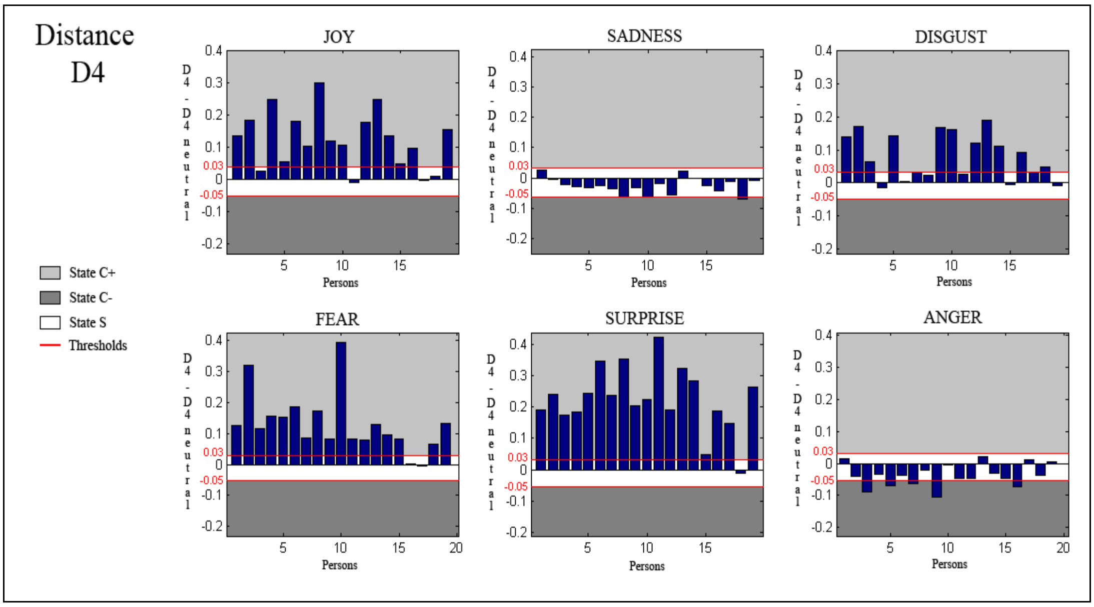


Fig. 4. Statistics results obtained for distance D4. Thresholds estimations are also shown.

	D_1	D_2	D_3	D_4	D_5	Wrinkles	Ang 1	Ang 2	W/H
Joy	C-	S/C-	C+	C+	C-	No	C+	S/C+/C-	S/C-
Surprise	S/C+	S/C+	S/C-	C+	S/C+	No	C-	C+	C-
Disgust	C-	C-	S/C+/C-	S/C+	S/C-	Yes	S/C+/C-	S/C+	S/C-
Anger	C-	C-	S/C-	S/C-	S/C+/C-	Yes	C+	C-	C+
Sadness	C-	S	S/C-	S	S/C+	No	S/C+/C-	S/C-	S/C+
Fear	S/C+	S/C+/C-	C-	C+	S/C+	No	C-	C+	C-
Neutral	S	S	S	S	S	No	S	S	S

Table 3. Theoretical table of the states taken by the different studied characteristics for each emotion, according to the results of the statistics obtained from the FG-NET database. The distances (D1,..D5) are those shown in Figure 2. Some features do not provide any information of interest for certain emotions (squares in gray) and in these cases they are not considered. The four last columns are explained in sections 3.2 and 3.3 and in Figure 6.

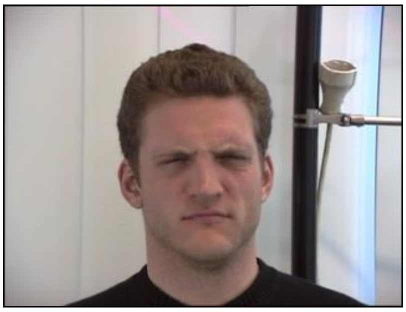


Fig. 5. Frame classified like “disgust” by the FG-NET database [12].

The obtained results are shown in the second column in Table 4. As it can be observed, the percentage of success obtained for the emotions “disgust”, “anger”, “sadness”, “fear” and “neutral” are acceptable (68.42-100%). The results for “disgust” and “neutral” are widely better than the obtained by Hammal [10]. Moreover our method is able to classify into one of the 7 universal basic emotions, while Hammal's only manages 4 emotional categories. Nevertheless, for “joy” and “surprise” the results are not very favourable. In fact, the algorithm tends to confuse “joy” with “disgust” and “surprise” with “fear”, which comes justified looking at Table 3, where it can be seen that a same combination of states of distances can be given for the mentioned pairs of emotions. In Table 4, as well as comparing our method with Hammal’s method [10], there is also a comparison with other recent works with the same experimental proposal [23,24,25]. Information about the type of classifier, the type of model (based on facial features or holistic model) and the characteristics of the different databases is included for each work. It is important to realize that the database used in our work is bigger than the used in the other ones (1500 images of 62 individuals of all races and genders), and therefore more universal.

Related to classification success, it is interesting to realize that human mechanisms for face detection are very robust, but this is not the case of those for face expressions interpretation. According to Bassili [14], a trained observer can correctly classify faces showing emotions with an average of 87%.

3.2 Addition of characteristics: information about the wrinkles in the nasal root

In order to improve the results obtained in “joy”, we introduce a new face parameter: the presence or absence of wrinkles in the nasal root, typical of the emotions “disgust” and “anger”. The new parameter is automatically detected by a real-time Gabor filter-based feature extraction program. This way, we will mark a difference between “joy” and “disgust”. The obtained success rates are shown in the third column in Table 4. We observe, as it was expected, a considerable increase in the rate of successes, especially for “joy” and “disgust”. However, the rates still continue being low for “sadness” and “surprise”, which makes us think about the necessity to add more characteristics to the method.

3.3 Addition of characteristics: information about the mouth shape

A key factor to analyze in the recognition of emotions is the mouth shape. For each one of the 7 basic emotions, its contour changes in many different ways. In our method, we have added the extra information about the mouth behaviour that is shown in Figure 6. Results are shown in the fourth column at Table 4. As it can be seen, the new information introduced a great improvement in our results. The importance of the mouth shape in the expression of emotions is thus confirmed.

		Our method	Our method + wrinkles	Our method + mouth shape	Method of Hammal et al. [10]	Method of Datcu & Rothkrantz [23]	Method of Wang et al. [24]	Method of Kuilenburg et al. [25]
Type of classifier		rule-based	rule-based	rule-based	rule-based	SVM	AdaBoost	Active Appearance Model
Type of model		facial features	facial features	facial features	facial features	facial features	facial features	holistic
Database		1500 frames, 62 subjects	1500 frames, 62 subjects	1500 frames, 62 subjects	630 frames, 8 subjects	474 frames	213 frames, 9 japanese females	980 high-quality facial images
Success rates	**Joy**	36.84%	100%	100%	87.26%	72.64%	93.50%	97.00%
	Surprise	57.89%	63.16%	63.16%	84.44%	83.80%	92.50%	85.00%
	Disgust	84.21%	94.74%	100%	51.20%	80.35%	93.25%	88.00%
	Anger	73.68%	94.74%	89.47%	not recognized	75.86%	92.50%	80.00%
	Sadness	68.42%	57.89%	94.74%	not recognized	82.79%	91.75%	85.00%
	Fear	78.95%	84.21%	89.47%	not recognized	84.70%	91.25%	93.00%
	Neutral	100%	100%	100%	88.00%	not recognized	91.50%	96.00%

Table 4. Classification rates of our method with the 5 distances; plus wrinkles in the nasal root; plus mouth shape information; and comparison with the rates obtained by Hammal [10] and other recent approaches [23, 24, 25].

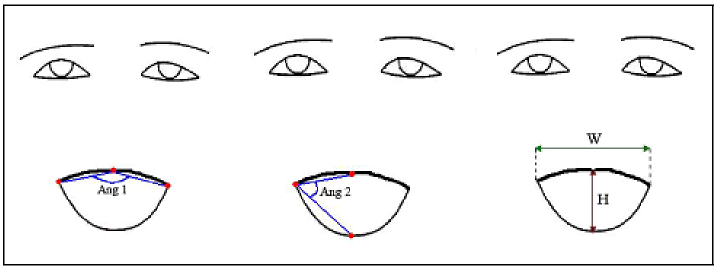


Fig. 6. Extra information added about the mouth shape.

3.4 Influence of race in the classification results

In this section we analyze the influence of the race in the studied face characteristics. To do it, we have used the JAFFE database (Japanese Female Facial Expression Database, [16]), that contains photographs of 10 Japanese women expressing the 6 basic Ekman's emotions (Fig. 7), and we have compared it with the databases used to implement our system (FG-NET and MMI) that mainly contain Caucasian individuals.

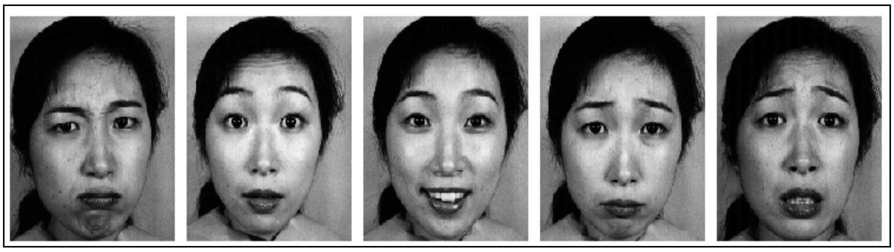


Figure 7: Examples of photographs of the JAFFE database [16].

Studying the obtained graphs, we can observe that the evolution of the different characteristics with respect to the neutral expression for the Asians is much smaller than for the Caucasian race (Fig. 8). This makes difficult the establishment of a criterion of classification for the system since, although the evolution of the parameters follows the same tendency that in the Caucasian race (states C^+, C^- and S), such small variations make difficult the establishment of the thresholds that allow to clearly discriminate among emotions. The limited accuracy of the facial points' automatic extraction algorithm also makes this task difficult as it can happen to be of the same order that the variation of the characteristics.

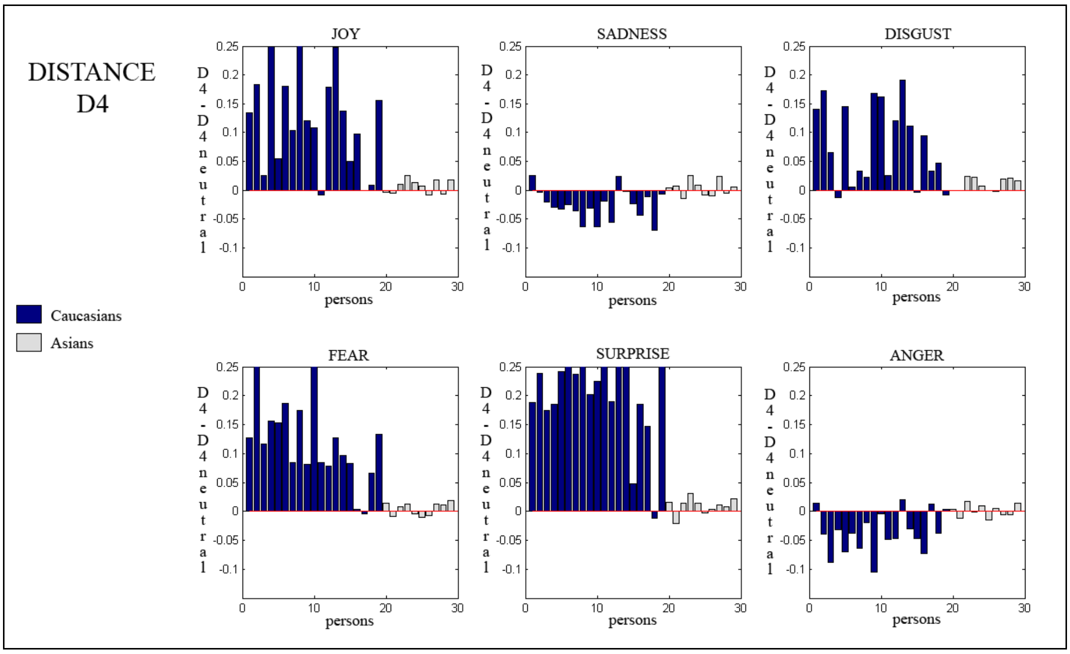


Figure 8: Comparison of the values taken by the distance D_4 for Caucasians and Asians.

The smaller expressiveness of the Asian race can come from several factors. On the one hand, from the characteristics of the mongoloid appearance (torn eyes, sunk nose, form of the cheekbone, etc.), that reduce flexibility to the face. On the other hand, from geographic, historical, cultural, educational and traditional factors, that also can play an essential role in the way of expressing emotions in the different races.

The indicated previously is not in discord with the Ekman's theory of Universality of emotions [9]. The emotions are universal, a Caucasian is perfectly able to recognize and to classify into an emotional category the expression shown by an Asian and vice versa, the only difference between them is the intensity of the shown emotions.

4 Application for real-time natural interfaces

4.1 Temporal information: analysing video sequences

After having tuned and validated the classification system with the static images, the use of the automatic feature extraction has enabled us to track video sequences of user's caught by a webcam. Psychological investigations argue that the timing of the facial expressions is a critical factor in the interpretation of expressions. In order to give temporary consistency to the system, a temporary window that contains the emotion detected by the system in each one of the 9 previous frames is created. A variation in the emotional state of the user is detected if in this window the same emotion is repeated at least 6 times and is different from the detected in the last emotional change.

The parameters corresponding to the neutral face are obtained calculating the average of the first frames of the video sequence, in which the user is supposed to be in the neutral state. For the rest of the frames, a classification takes place following the method explained in the previous sections.

4.2 Application: new input data for natural interfaces

To demonstrate the potential of our emotional tracking system, we have added it to Maxine [15], a general engine for real-time management of virtual scenarios and characters developed by the group. Maxine is a tool that has been created with the aim of making it easy the use of character-based interfaces in different application domains (educational tutors, virtual presenters, domotic helpers…). The general vision is that if a user's emotion could be recognized by computer, human interaction would become more natural, enjoyable and productive. The computer could offer help and assistance to a confused user or try to cheer up a frustrated user, and hence react in more appropriate ways.

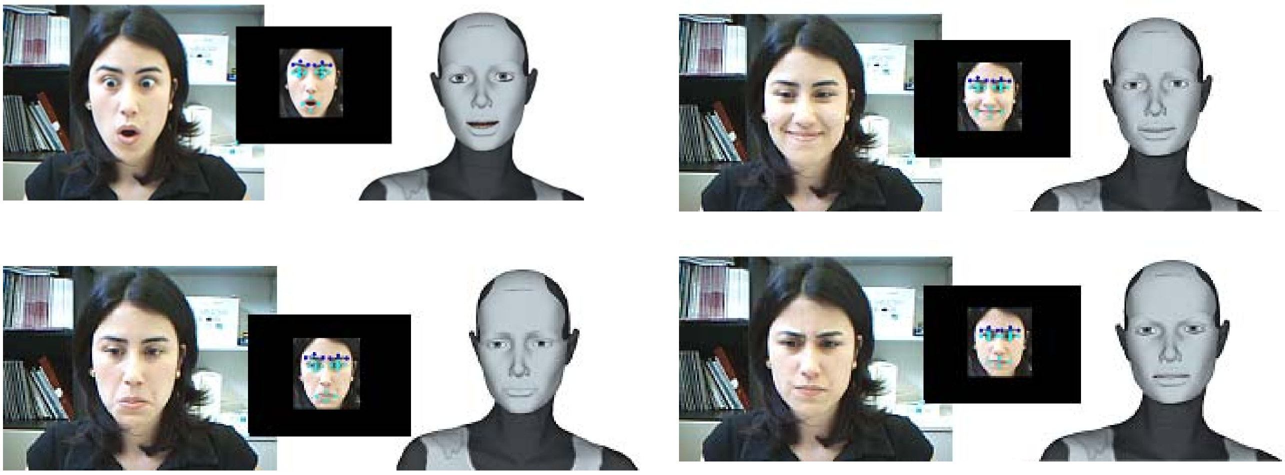


Fig. 9. Examples of the integrated real-time application: detection of surprise, joy, sadness, anger. For each example, images caught by the webcam, small images showing automatic features' tracking and synthesized facial expressions are shown. The animated character mimics the facial expression of the user.

The system presented here has been configured as a new multimodal input to the system. The system recognizes the emotion of the user and responds in an engaging way. The features extraction program [11] captures each facial frame and extracts the feature points which are sent to the emotion classifier. When an emotional change is detected, the output of the 7-emotion classifier constitutes an emotion code which is sent to Maxine's character.

In order to appreciate the power of our emotional classification method, we have make the virtual character's face mimic the emotional state of the user (Fig. 9), accommodating his/her facial animation and speech.

5 Conclusions and future work

We have presented a simple and effective method for the automatic classification of facial expressions. The introduction of several additional parameters barely increases the computational cost of the algorithm, given its simplicity, and produces very significant rates of improvement. In the future it is hoped to introduce new characteristics, in the form of face distances or angles (for example the angle formed by the eyebrows). The automatic features extraction program allows the introduction of dynamic information in the classification system, making it possible the study of the time evolution of the evaluated parameters, and the classification of user's emotions from live video.

To prove its usefulness and real-time operation, the system has been added to the Maxine system, an engine developed by the group for managing 3D virtual scenarios and characters to enrich user interaction in different application domains. For the moment, and as a first step, the emotional information has been used to accommodate facial animation and speech of the virtual character to the emotional state of the user. More sophisticated adaptive behaviour is now being explored. As it has been pointed out, recognition of emotional information is a key step toward giving computers the ability to interact more naturally and intelligently with people.

Acknowledgments

The authors wish to thank the Computer Graphics, Vision and Artificial Intelligence Group of the University of the Balearic Islands for providing us the real-time facial tracking module to test our classifier.

This work has been partially financed by the Spanish “Dirección General de Investigación” Nº TIN2004-07926 and TIN2007-63025, and by the Aragon Government through the WALQA Agreement Ref. 2004/04/86 and the CTPP02/2006 project.

References

[1] V. Bruce: “What the Human Face Tells the Human Mind: some Challenges for the Robot-Human Interface”, *Proc. Int’l workshop Robot and Human Comm.A.*, 44-51, 1992.

[2] G.J. Edwards, T.F. Cootes, C.J. Taylor, “Face Recognition Using Active Appearance Models”, *Proc. European Conf. Computer Vision*, Vol. 2, 581-695, 1998.

[3] H. Hong, H. Neven, C. von der Malsburg, “Online Facial Expression Recognition Based on Personalized Galleries”, *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 354-359, 1998.

[4] M.J. Lyons, J. Budynek, S. Akamatsu, “Automatic Classification of Single Facial Images”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, nº12, 1357-1362, 1999.

[5] Z. Zhang, M. Lyons, M. Schuster, S. Akamatsu, “Comparison between Geometry-Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron”, *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, 454-459, 1998.

[6] M. Wallace, A. Raouzaiou, N. Tsapatsoulis, S. Kollias, “Facial Expression Classification Based on MPEG-4 FAPs: The Use of Evidence and Prior Knowledge for Uncetainty Removal”, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Vol. 1, 51-54, 2004.

[7] M. Pantic, L.J.M. Rothkrantz, “Expert System for Automatic Analysis of Facial Expression”, *Image and Vision Computing J.*, Vol. 18, nº. 11, 881-905, 2000.

[8] M. Pantic, L.J.M. Rothkrantz: “Automatic Analysis of Facial Expressions: The State of the Art”, *Pattern Analysis and Machine Intelligence, IEEE Transactions*, Vol. 22, Issue 12, 1424–1445, 2000.

[9] P. Ekman: “Facial Expression, the Handbook of Cognition and Emotion”, John Wiley et Sons, 1999.

[10] Z. Hammal, L. Couvreur, A. Caplier, M. Rombaut: “Facial Expressions Recognition Based on the Belief Theory: Comparison with Diferent Classifiers”, *Proc. 13th International Conference on Image Analysis and Processing*, 2005.

[11] C. Manresa-Yee, J. Varona, F.J. Perales: “Towards hands-free interfaces based on real-time robust facial gesture recognition”. *AMDO’06, Lecture Notes In Computer Science*, nº4069, 504-513, 2006.

[12] http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html (Reviewed in February 2006).

[13] M. Pantic, M.F. Valstar, R. Rademaker, L. Maat: "Web-based Database for Facial Expression Analysis", *Proc. IEEE Int'l Conf. Multmedia and Expo (ICME'05)*, 2005.

[14] J.N. Bassili: “Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face”. *Journal of Personality and Social Psychology*, nº 37, 2049-2059, 1997.

[15] F. Seron, S. Baldassarri, E. Cerezo: “MaxinePPT: Using 3D Virtual Characters for Natural Interaction”, *Proc. WUCAmI’06: 2nd International Workshop on Ubiquitous Computing and Ambient Intelligence*, 241–250, 2006.

[16] M.J. Lyons, J. Budynek , S. Akamatsu: “Automatic Classification of Single Facial Images”. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, nº12, 1357-1362, 1999.

[17] Y. Duang-Duang, J. Lian-Wen, Y. Jun-Xun, Z. Li-Xin, H. Jian-Cheng: “Facial expression recognition with Pyramid Gabor Features and Complete Kernel Fisher Linear Discriminant Analysis”, *International Journal of Information Technology*, Vol. 11, nº 9, 91-100, 2005.

[18] M. Limin, D. Chelberg, M. Celenk: "Spatio-temporal modeling of facial expressions using Gabor-wavelets and hierarchical hidden Markov models", *IEEE International Conference on Image Processing*, Vol. 2, 57-60, 2005.

[19] I. Cohen, N. Sebe, A. Garg, L.S. Chen, T.S. Huang: "Facial expression recognition from video sequences: temporal and static modelling", *Computer Vision and Image Understanding*, Vol. 11, nº 1-2, 160-187, 2003.

[20] N. Esau, E. Wetzel, L. Kleinjohann, B. Keinjohann: "Real-time facial expression recognition using a fuzzy emotion model", *IEEE International Fuzzy Systems Conference FUZZ-IEEE 2007*, 1-6, 2007.

[21] S. Ioannou, A. Raouzaiou, V. Tzouvaras, T. Mailis, K. Karpouzis, S. Kollias: "Emotion recognition through facial expression analysis based on a neurofuzzy network", *Special Issue on Emotion: Understanding & Recognition*, Neural Networks, Elsevier, Vol. 18, Issue 4, 423-435, May 2005.

[22] X. Qinzhen, Z. Pinzheng, P. Wenjiang, Y. Luxi, H. Zhenya: "A Facial Expression Recognition Approach Based on Confusion-Crossed Support Vector Machine Tree", *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 309-312, 2006.

[23] D. Datcu, L.J.M. Rothkrantz: "Facial Expression Recognition in still pictures and videos using Active Appearance Models. A comparison approach.", *Proceedings of the 2007 international conference on Computer systems and technologies*, Vol. 285, nº 112, 2007.

[24] Y. Wang, H. Ai, B. Wu, C. Huang: "Real Time Facial Expression Recognition with Adaboost", Proceedings of the 17^{th} International Conference on Pattern Recognition ICPR 2004, Vol. 3, 926-929, 2004.

[25] H. vans Kuilenburg, M. Wiering, M. den Uyl: "A Model Based Method for Automatic Facial Expression Recognition", Proceedings of the 16th European Conference on Machine Learning, 194-205, October 2005.