Sensing Facial Emotions in a Continuous 2D Affective Space

Isabelle Hupont
Reference Centre on Interactive Television
Aragon Institute of Technology
Zaragoza, Spain
ihupont@ita.es

Eva Cerezo, Sandra Baldassarri
Computer Science and Systems Engineering Department
University of Zaragoza
Zaragoza, Spain
{ecerezo, sandra}@unizar.es

***Abstract*—The interpretation of user facial expressions is a very useful method for emotional sensing and it constitutes an indispensable part of affective Human Computer Interface designs. Facial expressions are often classified into one of several basic emotion categories. This categorical approach seems poor to treat faces with blended emotions, as well as to measure the intensity of a given emotion. This paper presents an effective system for facial emotional classification, where facial expressions are evaluated with a psychological 2-dimensional continuous affective approach. At its output, an expressional face is represented as a point in a 2D space characterized by evaluation and activation factors. The proposed system first starts with a classification method in discrete categories based on a novel combination of classifiers, that is subsequently mapped in a 2D space in order to be able to consider intermediate emotional states. The system has been tested with an extensive universal database and human assessment has been taken into consideration in the evaluation of results.**

***Keywords*—Kansei (sense/emotion) engineering, human factors, affective computing, facial expression analysis.**

I. INTRODUCTION

Human computer intelligent interaction is an emerging field aimed at providing natural ways for humans to use computers as aids. It is argued that for a computer to be able to interact with humans it needs to have the communication skills of humans. One of these skills is the affective aspect of communication [1].

The most expressive manner humans display emotions is through facial expressions. Facial expression is the most powerful, natural and direct way used by humans to communicate and understand each other's affective state and intentions [2]. Thus, the interpretation of facial expressions is the most common method used for emotional detection and forms an indispensable part of affective Human Computer Interface (HCI) designs.

The most long-standing way that facial affect has been described by psychologists is in terms of discrete categories, an approach that is rooted in the language of daily life. Facial expressions are often evaluated by classifying face images into the six universal emotions proposed by Ekman [3] which include "happiness", "sadness", "fear", "anger", "disgust" and "surprise". Examples of studies using this categorization are [4], [5]. The labeling scheme based on category is very intuitive and thus matches peoples' experience. This *categorical* approach, where emotions are a mere list of labels, fails however to describe the wide range of emotions that occur in daily communication settings and intrinsically ignore the intensity of an emotion. In this case, a small variation on face due to emotion may still be regarded as "neutral" face. There are a few tentative efforts to detect non-basic affective states from deliberately displayed facial expressions, including "fatigue" [6], and mental states such as "agreeing", "concentrating", "interested", "thinking", "confused", and "frustrated" [7], [8]. In any case, *categorical* approach presents a discrete list of emotions with no real link between them. It does not represent a *dimensional* space and has no algebra: every emotion must be studied and recognized independently.

To overcome the problems cited above, some researchers, such as Whissell [9] and Plutchik [10], prefer to view affective states not independent of one another but rather related to one another in a systematic manner. They consider emotions as a continuous 2D space whose dimensions are evaluation and activation. The evaluation dimension measures how a human feels, from positive to negative. The activation dimension measures whether humans are more or less likely to take some action under the emotional state, from active to passive. Besides *categorical* approach, *dimensional* approach is attractive because it provides a way of describing a wide range of emotional states and measuring the intensity of emotion. It is much more able to deal with non-discrete emotions and variations in emotional states over time, since in such cases changing from one universal emotion label to another would not make much sense in real life scenarios. However, in comparison with category-based description of affect, very few works have chosen a dimensional description level, and the few that do are more related to the design of synthetic faces [11], data processing [12] or psychological studies [13] than to emotion recognition. Moreover, in existing affective recognition works the problem is simplified to a two-class (positive vs negative and active vs passive) [14] or a four class (quadrants of 2D space) classification [15], thereby losing the descriptive potential of 2D space. Apart from seeking effective features that reflect affective factors, the main difficulty comes from the labeling of ground-truth data since there is any

This work has been partly financed by the Spanish DGICYT contract number TIN2007-63025, by the Government of Aragon: IAF Nº2008/0574 and CyT Nº2008/0486 agreements and "Plan Avanza" contract number TSI-020100-2009-828.

978-1-4244-6587-3/10/$25.00 É2010 IEEE

available public facial expression database that provides emotional annotations in terms of evaluation and activation dimensions.

Independently of the description level chosen to classify emotions (*categorical* or *dimensional*), a classification mechanism must be establish to categorize the facial posture shown in terms of the defined description level. In the literature, the facial expression analyzers that obtain the best success rates for emotional classification make use of neural networks, rule-based expert systems, Support Vector Machines or Bayesian nets based classifiers. In [16], an excellent state-of-the art summary is given of the various methods recently used in facial expression emotional recognition. However, the majority of those studies confine themselves to select only one type of classifier for emotional detection, or at the most compare different classifiers and then use that which provides the best results [5].

In this paper, an effective system for sensing facial emotions in a continuous 2D affective space is described. Its inputs are a set of carefully selected facial distances and angles that modelize the face in a simple way but without losing relevant facial expression information. The system starts with a classification method in discrete categories that is subsequently expanded in order to be able to work in a continuous emotional space and thus to consider intermediate emotional states. As regards the classification mechanism itself, the system intelligently combines the outputs of different classifiers simultaneously. In this way, the overall risk of making a poor selection with a given classifier for a given facial input is considerably reduced. The system is capable of analyzing any subject, male or female of any age and ethnicity, and has been validated considering human assessment.

The structure of the paper is the following: Section 2 describes the classification method in discrete categories. In Section 3 the step from the discrete perspective to the continuous emotional space is explained in detail and Section 4 comprises conclusion and a description of future work.

II. A Novel Method for Discrete Emotional Classification

In this section, an effective method is presented for the automatic classification of facial expressions into discrete emotional categories. The method is able to classify the user's emotion in terms of the six Ekman's universal emotions (plus "neutral"), giving a confidence value to each emotional category. Section A explains the selection and extraction process of the features serving as inputs to the system. Section B describes the criteria taken into account when selecting the various classifiers and how they are combined. Finally, the obtained results are presented in section C.

A. Selection and Extraction of Facial Inputs

Facial Action Coding System (FACS) [17] was developed by Ekman and Friesen to code facial expressions in which the individual muscular movements in the face are described by Action Units (AUs). This work inspired many researchers to analyze facial expressions by means of image and video processing, where by tracking of facial features and measuring a set of facial distances and angles, they attempt to classify different facial expressions. In particular, existing works demonstrate that a high emotional classification accuracy can be obtained by analyzing a small set of facial distances and angles. Examples are the work of Soyel and Demirel [18] that studies six 3D facial distances; the method proposed by Hammal et al. [4], that analyzes a set of five 2D facial distances; or the approach of Chang et al. [19], that measures twelve feature distances.

Following that methodology, the initial inputs of our classifiers were established in a set of distances and angles obtained from 20 characteristic facial points. In fact, the inputs are the variations of these angles and distances with respect to the "neutral" face. The chosen set of initial inputs compiles the distances and angles that have been proved to provide the best classification performance in existing works of the literature, such as the aforementioned. The points are obtained thanks to faceAPI [20], a commercial real-time facial feature tracking program that provides Cartesian facial 3D coordinates. It is able to track up to +/- 90 degrees of head rotation and is robust to occlusions, lighting conditions, presence of beard, glasses, etc. The initial set of parameters tested is shown in Fig. 1. In order to make the distance values consistent (independently of the scale of the image, the distance to the camera, etc.) and independent of the expression, all the distances are normalized with respect to the distance between the eyes. The choice of angles provides a size invariant classification and saves the effort of normalization.

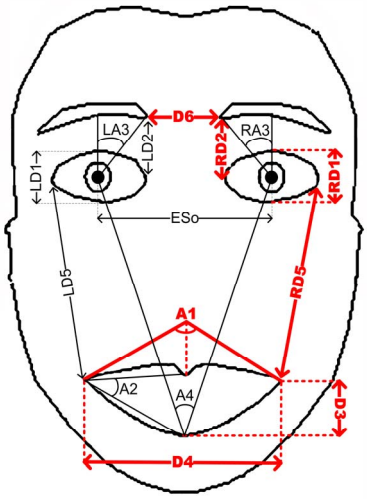

Figure 1. Facial parameters tested (in bold, the final selected parameters).

In order to determine the goodness and usefulness of the parameters, a study of the correlation between them was carried out using the data (distance and angle values) obtained from a set of training images. For this purpose, two different facial emotion databases were used: the FGNET database [21] that provides spontaneous (non-acted) video sequences of 19 different young Caucasian people, and the MMI Facial Expression Database [22] that holds 1280 acted videos of 43 different subjects from different races (Caucasian, Asian, South American and Arabic) and ages ranging from 19 to 62. Both databases show Ekman's six universal emotions plus the "neutral" one and provide expert annotations about the emotional apex frame of the video sequences. A new database

has been built for this work with a total of 1500 static frames selected from the apex of the video sequences from the FG-NET and MMI databases. It has been used as a training set in the correlation study and in the tuning of the classifiers.

A correlation-based feature selection technique [23] was carried out in order to identify the most influential parameters in the variable to predict (emotion) as well as to detect redundant and/or irrelevant features. Subsets of parameters that are highly correlated with the class while having low intercorrelation are preferred. In that way, from the initial set of parameters only the most significant ones were selected to work with: RD1, RD2, RD5, D3, D4, D6 and A1 (marked in bold in Fig. 1). This reduces the number of irrelevant, redundant and noisy inputs in the model and thus computational time, without losing relevant facial information.

B. Classifiers Selection and Novel Combination

In order to select the best classifiers, the Waikado Environment for Knowledge Analysis (Weka) tool was used [24]. This provides a collection of machine learning algorithms for data mining tasks. From this collection, five classifiers were selected after tuning and benchmarking: RIPPER, Multilayer Perceptron, SVM, Naive Bayes and C4.5. The selection was based on their widespread use as well as on the individual performance of their Weka implementation.

A 10-fold cross-validation test over the 1500 training images has been performed for each selected classifier. The success rates obtained for each classifier and each emotion are shown in the first five rows of Table I. As can be observed, each classifier is very reliable for detecting certain specific emotions but not so much for others. For example, the C4.5 is excellent at identifying "joy" (92.90% correct) but is only able to correctly detect "fear" on 59.30% of occasions, whereas Naive Bayes is way above the other classifiers for "fear" (85.20%), but is below the others in detecting "joy" (85.70%) or "surprise" (71.10%). Therefore, an intelligent combination of the five classifiers in such a way that the strong and weak points of each are taken into account appears as a good solution for developing a method with a high success rate.

TABLE I. SUCCESS RATES OBTAINED WITH A 10-FOLD CROSS-VALIDATION TEST OVER THE 1500 TRAINING IMAGES FOR EACH INDIVIDUAL CLASSIFIER AND EMOTION (FIRST FIVE ROWS) AND WHEN COMBINING THE FIVE CLASSIFIERS (SIXTH ROW).

	Disgust	**Joy**	**Anger**	**Fear**	**Sadness**	**Neutral**	**Surprise**
RIPPER	50.00%	85.70%	66.70%	48.10%	26.70%	80.00%	80.00%
SVM	76.50%	92.90%	55.60%	59.30%	40.00%	84.00%	82.20%
C4.5	58.80%	92.90%	66.70%	59.30%	30.00%	70.00%	73.30%
Naive Bayes	76.50%	85.70%	63.00%	85.20%	33.00%	86.00%	71.10%
Multilayer Perceptron	64.70%	92.90%	70.40%	63.00%	43.30%	86.00%	77.80%
Combination of classifiers	**94.12%**	**97.62%**	**81.48%**	**85.19%**	**66.67%**	**94.00%**	**95.56%**

The classifier combination chosen follows a weighted majority voting strategy. The voted weights are assigned depending on the performance of each classifier for each emotion. From each classifier, a confusion matrix formed by elements $P_{jk}(E_i)$, corresponding to the probability of having emotion i knowing that classifier j has detected emotion k, is obtained. The probability assigned to each emotion $P(E_i)$ is calculated as:

$$P(E_i) = \frac{P_{1k'}(E_i) + P_{2k''}(E_i) + \ldots + P_{5k^v}(E_i)}{5} \quad (1)$$

where: k', k'' ... k^v are the emotions detected by classifiers 1, 2 ... 5, respectively.

The assignment of the final output confidence value corresponding to each basic emotion is done following two steps:

1) Firstly, the confidence value $CV(E_i)$ is obtained by normalizing each $P(E_i)$ to a 0 through 1 scale:

$$CV(E_i) = \frac{P(E_i) - \min\{P(E_i)\}}{\max\{P(E_i)\} - \min\{P(E_i)\}} \quad (2)$$

where:

- $\min\{P(E_i)\}$ is the greatest $P(E_i)$ that can be obtained by combining the different $P_{jk}(E_i)$ verifying that $k \neq i$ for every classifier j. In other words, it is the highest probability that a given emotion can reach without ever being selected by any classifier.
- $\max\{P(E_i)\}$ is that obtained when combining the $P_{jk}(E_i)$ verifying that k=i for every classifier j. In other words, it is the probability that obtains a given emotion when selected by all the classifiers unanimously.

2) Secondly, a rule is established over the obtained confidence values in order to detect and eliminate emotional incompatibilities. The rule is based on the work of Plutchik [10], who assigned "emotional orientation" values to a series of affect words. For example, two similar terms (like "joyful" and "cheerful") have very close emotional orientation values while two antonymous words (like "joyful" and "sad") have very distant values, in which case Plutchik speaks of "emotional incompatibility". The rule to apply is the following: if emotional incompatibility is detected, i.e. two non-null incompatible emotions exist simultaneously, that chosen will be the one with the closer emotional orientation to the rest of the non-null detected emotions. For example, if "joy", "sadness" and "disgust" coexist, "joy" is assigned zero since "disgust" and "sadness" are emotionally closer according to Plutchik.

C. Results

The results obtained when applying the strategy explained in the previous section to combine the scores of the five classifiers with a 10-fold cross-validation test are shown in sixth row of Table I. As can be observed, the success rates for the "neutral", "joy", "disgust", "surprise", "disgust" and "fear" emotions are very high (81.48%-97.62%). The lowest result of our classification is for "sadness", which is confused with the "neutral" emotion on 20% of occasions, due to the similarity of their facial expressions. Confusion between this pair of emotions occurs frequently in the literature and for this reason many works do not consider "sadness". Nevertheless, the results can be considered positive as emotions with distant

"emotional orientation" values (such as "disgust" and "joy" or "neutral" and "surprise") are confused on less than 2.5% of occasions and incompatible emotions (such as "sadness" and "joy" or "fear" and "anger") are never confused. Table II shows the confusion matrix obtained after the combination of the five classifiers.

TABLE II. CONFUSION MATRIX OBTAINED COMBINING THE FIVE CLASSIFIERS.

Emotion --> is classified as	Disgust	Joy	Anger	Fear	Sadness	Neutral	Surprise
Disgust	**94,12%**	0,00%	2,94%	2,94%	0,00%	0,00%	0,00%
Joy	2,38%	**97,62%**	0,00%	0,00%	0,00%	0,00%	0,00%
Anger	7,41%	0,00%	**81,48%**	0,00%	7,41%	3,70%	0,00%
Fear	3,70%	0,00%	0,00%	**85,19%**	3,70%	0,00%	7,41%
Sadness	6,67%	0,00%	6,67%	0,00%	**66,67%**	20,00%	0,00%
Neutral	0,00%	0,00%	2,00%	2,00%	2,00%	**94,00%**	0,00%
Surprise	0,00%	0,00%	0,00%	2,22%	0,00%	2,22%	**95,56%**

III. A 2D EMOTIONAL SPACE FOR THE EXTRACTION OF CONTINOUS EMOTIONAL INFORMATION

As discussed in the introduction, the use of a discrete set of emotions (labels) for emotional classification has important limitations. To avoid these limitations and enrich the emotional output information from the system in terms of intermediate emotions, use has been made of one of the most influential evaluation-activation 2D models in the field of psychology: that proposed by Whissell [9]. Thanks to this, and following the methodology explained in section A, the final output of the system will be the (x,y) coordinates in the activation-evaluation space of the analyzed facial expression. The results of emotional classification obtained in the 2D space are analyzed in detail in section B taking human assessment into account.

A. Emotional Mapping to a Continuous Affective Space

In her study, Whissell assigns a pair of values <evaluation, activation> to each of the approximately 9000 carefully selected affective words that make up her "Dictionary of Affect in Language" [9]. Fig. 2 shows the position of some of these words in the evaluation-activation space. The next step is to build an emotional mapping so that an expressional face image can be represented as a point on this plane whose coordinates (x,y) characterize the emotion property of that face.

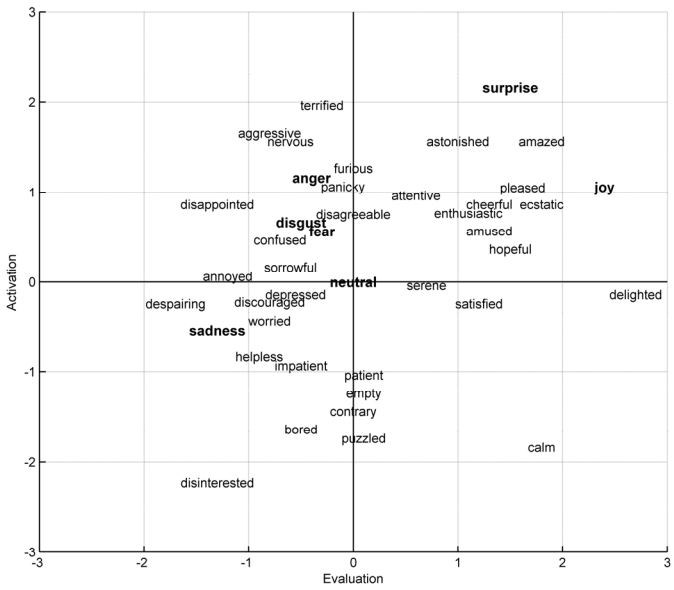

Figure 2. Simplified Whissell's evaluation-activation space.

It can be seen that the emotion-related words corresponding to each one of Ekman's six emotions have a specific location (x_i, y_i) in the Whissell space (in bold in Fig. 2). Thanks to this, the output information of the classifiers (confidence value of the facial expression to each emotional category) can be mapped in the space. This emotional mapping is carried out considering each of Ekman's six basic emotions plus "neutral" as 2D weighted points in the evaluation-activation space. The weights are assigned depending on the confidence value $CV(E_i)$ obtained for each emotion. The final (x,y) coordinates of a given image are calculated as the centre of mass of the seven weighted points in the Whissell space following (3). In this way the output of the system is enriched with a larger number of intermediate emotional states.

$$x = \frac{\sum_{i=1}^{7} x_i CV(E_i)}{\sum_{i=1}^{7} CV(E_i)} \quad and \quad y = \frac{\sum_{i=1}^{7} y_i CV(E_i)}{\sum_{i=1}^{7} CV(E_i)} \tag{3}$$

B. Evaluation of Results with Human Assessment

The method described in the previous section has been put into practice with the outputs of the classification system when applied to the database facial expressions images. In Fig. 3 the general location of all classified images is plotted (markers size is proportional to the percentage of images situated at the same location). Fig. 4 shows several images with their nearest label in the Whissell space.

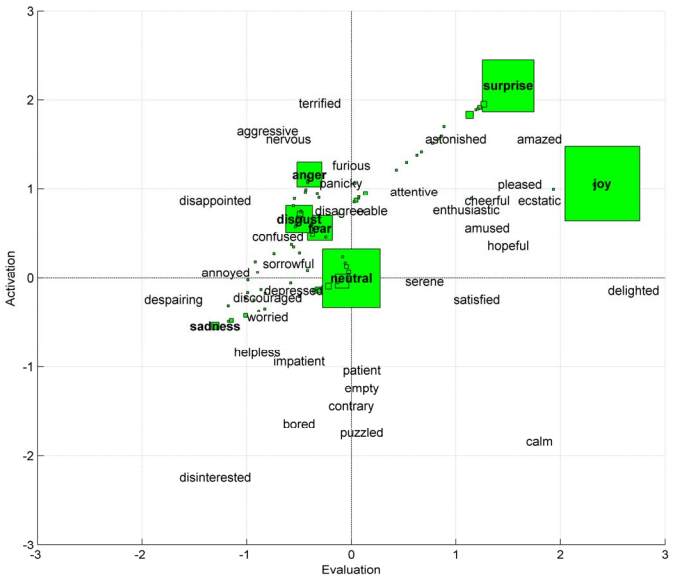

Figure 3. Location of the different images of the database in the Whissell space, according to the method explained in section III.A (marker size is proportional to the percentage of images situated at the same location).

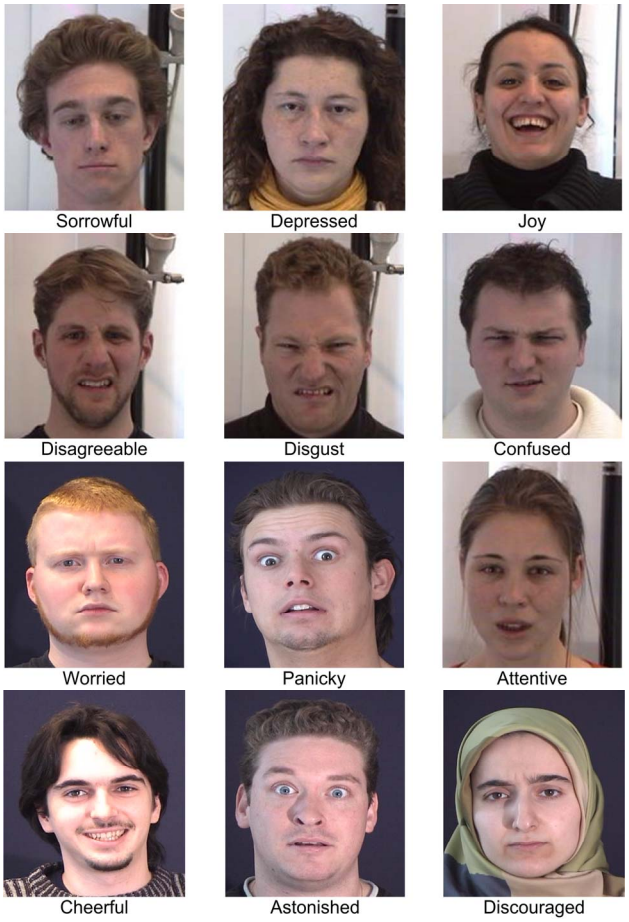

Figure 4. Examples of images from the database with their nearest label in the Whissell space, according to the method described in section III.A.

The database used in this work provides images labelled with one of the six Ekman universal emotions plus "neutral", but there is no a-priori known information about their location in the Whissell 2D space. In order to evaluate the system results, there is a need to establish the region in the Whissell space where each image can be considered to be correctly located. For this purpose, a total of 43 persons participated in one or more evaluation sessions (50 images per session). In the sessions they were told to locate a set of images of the database in the Whissell space (as shown in Fig. 2, with some reference labels). As result, each one of the frames was located in terms of evaluation-activation by 16 different persons.

The collected evaluation data have been used to define an ellipsoidal region where each image is considered to be correctly located. The algorithm used to compute the shape of the region is based on Minimum Volume Ellipsoids (MVE). MVE looks for the ellipsoid with the smallest volume that covers a set of data points. Although there are several ways to compute the shape of a set of data points (e.g. using a convex hull, rectangle, etc.), we chose the MVE because of the fact that real-world data often exhibits a mixture of Gaussian distributions, which have equi-density contours in the shape of ellipsoids. First, the collected data are filtered in order to remove outliers: a point is considered an outlier if its coordinate values (in both dimensions) are greater than the mean plus three times the standard deviation. Then, the MVE is calculated following the algorithm described by Kumar and Yildrim [25]. The MVEs obtained are used for evaluating results at four different levels:

1) *Ellipse criteria.* If the point detected by the system (2D coordinates in the Whissell space) is inside the defined ellipse, it is considered a success; otherwise it is a failure.

2) *Quadrant criteria.* The output is considered to be correctly located if it is in the same quadrant of the Whissell space as the ellipse centre.

3) *Evaluation axis criteria.* The system output is a success if situated in the same semi-axis (positive or negative) of the evaluation axis as the ellipse centre. This information is especially useful for extracting the positive or negative polarity of the shown facial expression.

4) *Activation axis criteria.* The same criteria projected to the activation axis. This information is relevant for measuring whether the user is more or less likely to take an action under the emotional state.

The results obtained following the different evaluation strategies are presented in Table III.

TABLE III. RESULTS OBTAINED ACCORDING TO DIFFERENT EVALUATION CRITERIA.

	Ellipse criteria	Quadrant criteria	Evaluation axis criteria	Activation axis criteria
Success rate	73.73%	87.45%	94.12%	92.94%

As can be seen, the success rate is 73.73% in the most restrictive case, i.e. when the output of the system is considered to be correctly located when inside the ellipse. It rises to 94.12% when considering the evaluation axis criteria.

Objectively speaking, these results are very good, especially when, according to Bassili [26], a trained observer can correctly classify facial emotions with an average of 87%. However, they are difficult to compare with other emotional classification studies that can be found in literature, given that either such studies do not recognize emotions in evaluation-activation terms, or they have not been tested under common experimental conditions (e.g. different databases or evaluation strategies are used).

IV. Conclusion and Future Work

This paper describes an effective method for facial emotional classification. The inputs to the system are a set of facial parameters (angles and distances between characteristic points of the face) that enable the face to be modeled in a computationally simple way without losing relevant information about the facial expression. The system combines in a novel manner the five most commonly used classifiers in the literature using a weighted majority voting strategy, obtaining at the output a confidence value of the facial expression to each of Ekman's six emotions (plus "neutral"). This emotional information is mapped on to Whissell's 2D evaluation-activation space with the aim of obtaining the location (coordinates) of the input facial expression in the space. The final output of the system does not, therefore, simply provide a classification in terms of a set of emotionally discrete labels, but goes further by extending the emotional information over an infinite range of intermediate emotions.

The main distinguishing feature of our work compared to others that use the evaluation-activation space for emotional classification is that the system output provides the exact location (coordinates) of facial expression in 2D space. Other works confine themselves to providing information about its polarity (positive/negative or active/passive) or the quadrant of space to which the image belongs. Another noteworthy feature of the work is that it has been tested with an extensive database of 1500 images showing individuals of different races and gender, giving universal results with very promising levels of correctness.

The recent focus on research area of affective computing lies on sensing emotions from multiple modalities, since natural human-human interaction is multimodal: people communicate through speech and use body language (posture, facial expressions, gaze) to express emotion, mood, attitude, and attention. A main question related to multimodality that still remains unsolved is how to fuse the information coming from different channels (audio, video, etc.). All available multimodal recognizers have designed and/or used ad-hoc solutions for fusing information coming from multiple modalities but cannot accept new modalities without re-defining and re-training the whole system. The use of a continuous emotional space in the way described in this paper opens the door to the fusion of different modules coming from different channels in a simple and scalable fashion. In fact, we are currently considering the integration of new multimodal emotional recognition input modules to the system (user's speech, gestures, gaze, mouse-clicks, keyboard use) making use of the Whissell space.

In a future it is also hoped to expand the method to pass from the analysis of still images to video sequences. Thanks to the use of the Whissell 2D continuous space, an emotional facial video sequence can be viewed as a point (corresponding to the location of a particular affective state in time *t*) moving through this space over time. In that way, the different positions taken by the point over time can be related mathematically and modeled to make the system more robust and consistent. The study of video sequences will open the door to analyze more samples to validate the system in more natural settings (e.g. movies, TV interviews, etc).

Acknowledgment

The authors wish to thank Dr. Cynthia Whissell for her explanations and kindness, Francisco Cruz for his collaboration in this work, and also all the participants in the evaluation sessions.

References

[1] H. Boukricha, C. Becker, and I. Wachsmuth, "Simulating empathy for the virtual human Max," Proceedings 2nd International Workshop on Emotion and Computing in conj. with the German Conference on Artificial Intelligence (KI2007), 2007, pp. 22–27.

[2] D. Keltner and P. Ekman, "Facial expression of emotion," Handbook of emotions, vol. 2, 2000, pp. 236–249.

[3] P. Ekman, T. Dalgleish, and M. Power, "Handbook of cognition and emotion," Chihester, UK: Wiley, 1999.

[4] Z. Hammal, A. Caplier, and M. Rombaut, "Belief theory applied to facial expressions classification," Pattern Recognition and Image Analysis, 2005, pp. 183-191.

[5] G. Littlewort, M.S. Bartlett, I. Fasel, J. Susskind, and J. Movellan, "Dynamics of facial expression extracted automatically from video," Image and Vision Computing, vol. 24, 2006, pp. 615–625.

[6] Q. Ji, P. Lan, and C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 36, 2006, pp. 862–875.

[7] A. Kapoor, W. Burleson, and R.W. Picard, "Automatic prediction of frustration," International Journal of Human-Computer Studies, vol. 65, 2007, pp. 724–736.

[8] M. Yeasin, B. Bullot, and R. Sharma, "Recognition of facial expressions and measurement of levels of interest from video," IEEE Transactions on Multimedia, vol. 8, 2006, pp. 500–508.

[9] C.M. Whissell, "The dictionary of affect in language," Emotion: Theory, Research and Experience, vol. 4, The Measurement of Emotions, New York: Academic, 1989.

[10] R. Plutchik, "Emotion: a psychoevolutionary synthesis," New York: Harper & Row, 1980.

[11] N. Stoiber, R. Seguier, and G. Breton, "Automatic design of a control interface for a synthetic face," Proceedings of the 13th International Conference on Intelligent User Interfaces, 2009, pp. 207–216.

[12] Y. Du, W. Bi, T. Wang, Y. Zhang, and H. Ai, "Distributing expressional faces in 2-D emotional space," Proceedings of the 6th ACM International Conference on Image and Video Retrieval, 2007, pp. 395–400.

[13] F. Gosselin and P.G. Schyns, "Bubbles: A technique to reveal the use of information in recognition tasks," Vision Research, vol. 41, 2001, pp. 2261–2271.

[14] N. Fragopanagos and J.G. Taylor, "Emotion recognition in human–computer interaction," Neural Networks, vol. 18, 2005, pp. 389–405.

[15] G. Garidakis, L. Malatesta, L. Kessous, N. Amir, A. Paouzaiou, and K. Karpouzis, "Modeling naturalistic affective states via facial and vocal

expression recognition," Int. Conf. on Multimodal Interfaces, 2006, pp. 146–154.

[16] Z. Zeng, M. Pantic, G.I. Roisman, and T.S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, pp. 39–58.

[17] P. Ekman, W.V. Friesen, and J.C. Hager, "Facial action coding system," Research Nexus eBook, 2002.

[18] H. Soyel and H. Demirel, "Facial expression recognition using 3D facial feature distances," Lecture Notes in Computer Science, vol. 4633, 2007, pp. 831–383.

[19] C.Y. Chang, J.S. Tsai, C.J. Wang, and P.C. Chung, "Emotion recognition with consideration of facial expression and physiological signals," Proceedings of the 6th Annual IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, 2009, pp. 278–283.

[20] Face API technical specifications brochure. Available: http://www.seeingmachines.com/pdfs/brochures/faceAPI-Brochure.pdf

[21] F. Wallhoff, "Facial expressions and emotion database," Technische Universität München, 2006. Available: http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html

[22] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," IEEE International Conference on Multimedia and Expo, 2005, pp. 317–321.

[23] M.A. Hall, "Correlation-based feature selection for machine learning," Hamilton, New Zealand, 1998.

[24] I. Witten and E. Frank, "Data Mining: practical machine learning tools and techniques," 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[25] P. Kumar and E.A. Yildirim, "Minimum-volume enclosing ellipsoids and core sets," Journal of Optimization Theory and Applications, vol. 126, 2005, pp. 1–21.

[26] 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, vol. 37, 1979, pp. 2049–2058.