Improving the Classification of Emotions by Studying Facial Feature

Y. Siwar, E. Ridha, J. Olfa, and Z. Mourad

Abstract—**This work aims to recognize the six basic emotions using facial expression and improve the classification in term of time and space memory.**

We started with the idea that emotions can be absolutely distinctive and from a single feature we can recognize emotion and therefore we save time in learning data.

We also noted that the similarity between the characteristics of two emotions can cause a lot of errors.

So we concluded that it is necessary to specify the characteristics for each emotion to improve classification.

The study of various facial features (eyes, eyebrows and mouth) of each emotion has helped us to find a new use of face features and ultimately lead to a faster classification.

Index Terms—**Emotion, recognition of emotions, emotion analysis, features extraction, facial expressions, face detection.**

I. Introduction

The automatically emotion recognition is a research topic that has growing interest in computer science. However, this is not an easy task. The act of recognizing is complex with a very diverse events and depends on many factoring (social, cultural and personality of the speaker).

The face is very important in communication of a human. From a physiological [1] point of view 55% of the communicative message is transferred by facial expressions.

So, facial expression is the most expressive part of a human. (Joy, surprise, disgust, sadness, anger and fear) are six universal emotions, which could be associated to 6 facial expressions by Ekman and Friesen [2]. They studied facial expressions in different cultures, including preliterate cultures, and found much commonality in the expression and recognition of emotions on the face. To recognize emotion from the face, it usually relies on the eyes, eyebrows and mouth [3] but these characteristics are sometimes similar and sometimes different between emotions. So using all features of face is not a good idea because it causes a waste of time, waste of space memory and more errors in classification.

In this work we have presented a new way to use the features of the face to improve the classification results in terms of time and memory space, I also present the steps that we followed to recognize emotions.

This paper is divided like this: at the first we will present the three steps of emotion recognizing which is: the face detection, the extraction of feature and the classification.

After, we will finish this article with results that we obtained using our approach.

II. Face Detection

The problem of face detection has been treated by several methods and techniques: we can mention the work of S. McKenna [4] who employs the color of skin by using simple rules to have a fast detection of face.

However, their main problem is the difficulty of choosing an appropriate color space and adequate decision rules that provide a high recognition.

Turk & Pentland [5] are the first introduced a method for face recognition based on 'eigenfaces'. They extract relevant information from a face image and encode it efficiently to compare to other. Relevant information is defined mathematically, instead of on human terms (such as lips, position of the eyebrows etc.), which is seen as a safe bet given the poor understanding of what the human neural system considers relevant in a face image. The idea is to capture the variation among a number of face images and use it to compare individual images.

More ever, there are the methods which were based on machine learning techniques: we can mention the work of H. Rowley [6]. For this, a set of learning data is divided on two parts: the first part present the images of the faces and the second part content the images without faces. So the problem becomes a problem of the classification. These are the most efficient algorithms but the learning step takes a long time and the number of images in the database can take a lot space of memory.

Fig. 1. Face detection by Viola and Jones.

Besides, we find another approach which is "Viola and Jones". This method detect all type of objects in a digital image that was been proposed by researchers [7]. It uses a complex algorithm of classification. It principle is to eliminate the objects that are not compatible with the intended object using the elementary classifiers. The method of Viola

and Jones combine several techniques that are already invented by researchers such as the integral image and the cascade classification. It is one of the best known and most widely used methods, in particular for the face detection. It detect efficiently and in real time faces in a picture.

As shown in Fig. 1, we used this method in this work.

III. FEATURE EXTRACTION

The methods of face detection are used also to extract the facial features (eyes, nose, and mouth).

Furthermore these two steps can be run simultaneously. This depends mainly on the nature of the application, the size of the training set, and shooting conditions (noise, dimming).

Extracting feature revolves around two ideas: the first is to extract the forms or texture of features (the eyes, mouth and eyebrows) and the second idea focus on using the muscle of face so we extract the distances.

For example we find Geometrical methods that use geometric shapes (circle, parabola, line, ellipse), such as the work of C.Loconsole [8] who use 8 ellipses and 3 lines to extract feature from face. These methods are more robust for the variation of face position, scale, size, and head orientation. However it is expensive and their performance is absolutely linked to the performance of face detection.

There are also the methods that use models defined a priori of the forms that we want to segment such as the method of Malciu [9]. These models are characterized by their ability to be deformable so that they adapt to the contours of the object.

Malciu defined a system based springs that connects corners of the feature of the face. His method is based on a function of the external energy to maintain uniformity between the image and the geometric model. He used the simplex method to minimize the total energy function and therefore have an optimal deformable model.

A local binary pattern (LBP) [10] is another method of feature extraction based on texture. It is robust to monotonic gray-scale changes caused, like the illumination variations. Also, it is simple because it can analyze images in real-time.

In addition, the most efficient method is the hybrid approach [11] that uses both geometric and texture methods. These approaches are excellent, but its takes a lot of time.

Transferable Belief Model [12], [13] is an approach that locates a set of point feature on the face, and then it calculates 21 distances whose:

- 7 distances: To present the action of the eyebrows.
- 2 distances: To present the action of the eyes.
- 2 distances: To present the action of the nose.
- 10 distances: To present the action of the mouth.

From the values of the distances, we can know the emotion of the input image.

This method locates well the facial features regardless of lighting conditions, color of skin and accessories (glasses, mustache). But his problem is that the images must be in front.
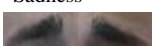
IV. CLASSIFICATION

A. Before Classification

They are a lot of psychologists who worked on the description of emotions. The most famous psychologists are Ekman and Friesen [14] who defined the FACS system (Facial Action coding system).They present a system of manual coding of facial expressions. Their work enables them to break down all visible facial movements in terms of 46 Action Units are a set of actions that corresponds to muscle movement such as raising lower lips, blinking, Biting lip, blow. Any observed of facial expression can be represented in the form of a combination of Shares Unitarians.

In this study we used the description of psychologies like Ekman to define and to compare the 6 basic emotions. After that, we constructed 15 tables that compare emotions pair wise using the features, the table below is an example of comparison between sadness and joy:

TABLE I: COMPARISON BETWEEN JOY/SADNESS

| Joy/Sadness	mouth	eyes	eyebrow
Example of an image	Joy Sadness	Joy Sadness	Joy Sadness
Description	+difference : • Sadness: The corners of the mouth tend downwards and the mouth is closed. • Joy: Corners of the mouth rise slightly with the possibility of appearance of the teeth.	+difference : • Sadness: Dropping upper eyelid • Joy: The "crow's feet" appear in the corners of the eyes.	+difference : • Sadness: The outer edges of the eyebrows don't move. The inner edges rise. • Joy: Eyebrows are neutral.
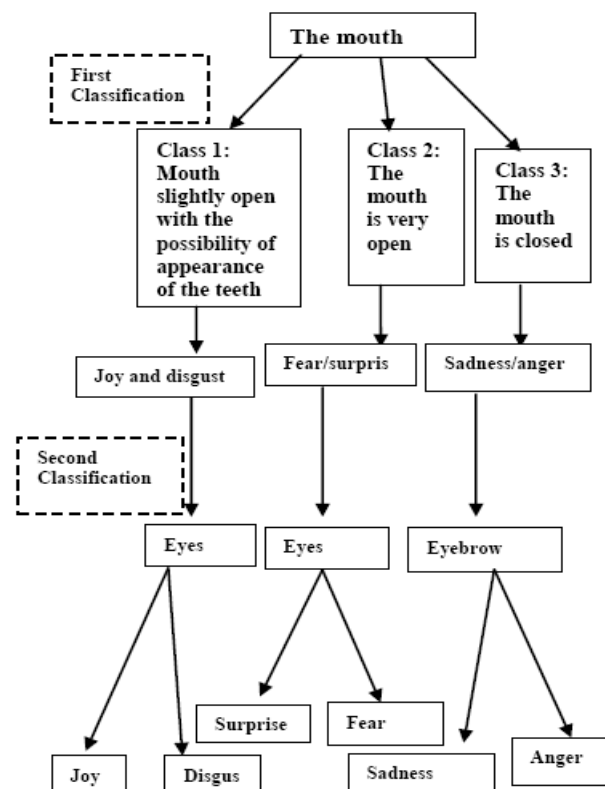
Fig. 2. Hierarchy used for classification.

From Table I we note that both sadness and happiness emotion are completely different. So from any feature we can distinguish between them.

Following these comparisons we have chosen the hierarchy defined in Fig. 2: It has two stages of classification:

The first classification is related to the mouth which will be classified according to three classes like this:

Class 1: Mouth is slightly open with the possibility of appearance of the teeth. It is the case of the happiness or disgust.

Class 2: The mouth is very open; it's the case of surprise and fear.

Class 3: The mouth is closed if emotion is sadness or anger.

In the second classification we used either the eyes or the eyebrows in order to classify the different emotions.

B. Classification

To recognize emotions, it is necessary to pass through the step of classifying. This phase uses the information extracted from the step of feature extraction to recognize what is the category of emotions (happiness, anger, sadness, etc.). Classification phase generally has two sub steps that are:

- The learning stage: They use a set of data that describes every emotion to estimate the description of the observation space. From this step we create a model of characteristic of each class.
- The test step: it is a step that evaluates the system. it uses a set of images to rate the performance of the classification system.

There are many algorithms for classification like neural networks [15] which are presented as systems of interconnected "neurons", its can compute values from inputs. A neural network is recognized by a set of input neurons that may be activated by the pixels of an input image. The actuations are passed to other neurons when the phase of weighted and transformed by a function is finished.
This action is repeated until finally, an output neuron is functioned. This identifies the emotion that was read.

KNN (K nearest neighbors) [16] is a simple algorithm. It selects k nearest observations. Therefore the decision is in favor of the majority class represented by k nearest neighbors.

The method of support vector machines (SVM) [17], [18] is an approach of classification by supervised learning. This method relies on the existence of a linear classifier in a suitable space It considers a hyper plane to separates the classes, ensuring that the margin between classes is maximal. This algorithm is used in this work.

V. EVALUATION

In this work, we used the data base of Cohn Kanade [12] to learning and testing our approach.

67% of data employed for learning and 33% of data for testing.

Using the hierarchy defined in Fig. 1 with SVM algorithm we got the confusion matrix as shown in Table II.

This table shows that the algorithm gives a very good result mainly in the case of joy, surprise and sadness. The average of recognition of six basics emotions by using our hierarchy is equal to 87%.

TABLE II: CONFUSION MATRIX OF OUR APPROACH USING SVM

Emotion	Joy	Fear	Sad	Anger	Disgust	Surprise
Joy	95%	0	0	0	5%	0
Fear	0	77%	5%	0	6%	12%
sad	0	0	89%	11%	0	0
Anger	0	0	17%	83%	0	0
Disgust	12%	0	5%	0	83%	0
Surprise	5%	0	0	0	0	95%

We also calculate the time and size of data in case of use of all the features at the same time.

The results are defined in Table III, which shows that we have reduced the time up to 1 or 2 seconds; also the size of data with our approach is 12 times lower than the classic approach.

TABLE III: OUR APPROACH COMPARED TO THE OTHER APPROACH THAT USES ALL THE FEATURES FOR CLASSIFICATION

	The approach that use all characteristics	Our approach
Time in learning (seconds)	Between 40 And 42	Between 1 and 2

VI. CONCLUSION

In conclusion, we have presented in this paper an application for the recognition of six basic emotions (happiness, disgust, anger, fear, surprise and sadness).
Also, we have set a target is the rapidity of recognition.

For the recognition of facial expression, we go through three main stages:
The first is the detection of face.
The second is the extraction of features that are (eyes, mouth and eyebrows).

The Flanker is the classification; primary we make a comparative study of different features of the face to find it a new way of using facial features for classifying emotion. This new way has improved classification in terms of time and space in memory.

ACKNOWLEDGMENT

REFERENCES

[1] A. Mehrabian, "Communication without words," *Psychology Today*, vol. 4, pp. 53-56, 1968.
[2] P. Ekman, "Facial expression," *The Handbook of Cognition and Emotion*, 1999.
[3] J. N. Bassili. "Facial motion in the perception of faces and of emotional expression," *Journal of Experimental Psychology*, vol. 4, pp. 373-379, 1978.
[4] S. McKenna, S. Gong, and Y. Raja, "Modeling facial color and identity with Gaussian mixtures," *Pattern Recognition*, 1998.
[5] Y. L. Tian, T. Kanade, and J. Cohn, "Evaluation of gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 229-234.
[6] H. Rowley, S. Baluja and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1998.
[7] P. Viola and M. Jones, "Robust real-time face detection," *Computer Vision and Pattern Recognition*, 2001.

[8] C. Loconsole1, C. RunaMiranda, G. Augusto, A. Frisoli1, and V. Orvalho, "Real-time emotion recognition: A novel method for geometrical facial features extraction," presented at International Conf. on Computer Vision Theory and Applications, January, 2014.

[9] M. Malciu and F. Preteux, "Mpeg-4 compliant tracking of facial features in video sequences," presented at International Conference on Augmented, Virtual Environments and 3D, 2001.

[10] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," *Image and Vision Computing*, 2009.

[11] A. Youssif, A. Aliaa, A. Wesam, and A. Asker, "Automatic facial expression recognition system based on geometric and appearance features," *Computer and Information Science*, 2011.

[12] Z. Hammal, L. Couvreur, A. Caplier, and M. Rombaut, "Facial expression recognition based on the belief theory: Comparison with different classifiers," presented at International Conference on Image Analysis and Processing, Cagliari, Italy, September, 2005.

[13] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," presented at IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[14] P. Ekman and W. Friesen, "Facial action coding system: A technique for the measurement of facial movement," *Consulting Psychologists Press*, 1978.

[15] B. Wang, "Facial expression recognition using neural networks," Technical Report, 1999.

[16] A. Sohail and P. Bhattacharya, "Classification of facial expressions using K-nearest neighbor classifier," *Computer Vision/Computer Graphics Collaboration Techniques*, 2007.

[17] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines," presented at IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[18] B. Boser, I. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. ACM Workshop on Computational Learning Theory*, 1992, pp. 144-152.

Siwar Yahia was born in Gabes, Tunisia in 1990. She obtained the bachelor degree in computer science in 2009. She received the master degree in computer science and multimedia and license degree in computer science from Higher Institute of Computer Science and Multimedia of Gabes (ISIMG), Gabes respectively in 2015 and 2012. She is interested in research related to computer technology, computer vision and image analysis.

Ridha Ejbali received the Ph.D. degree in computer engineering, the master degree and computer engineer degree from the National Engineering School of Sfax Tunisia (ENIS) respectively in 2012, 2006 and 2004.

He was an assistant technologist at the Higher Institute of Technological Studies, Kebili Tunisia since 2005. He joined the Faculty of Sciences of Gabes Tunisia (FSG) where he is an assistant in the Department of Computer Sciences since 2012. His research area is now in pattern recognition and machine learning using Wavelets and Wavelet networks theories.

Olfa Jemai received her B.S. degree in computer science from the National School of Computer Sciences of Tunis (ENSI) in 1999. She obtained her M.S. and PhD degrees in computer engineering from the National Engineering School of Sfax (ENIS), in 2004 and 2010, respectively. She spent seven years as a contractual assistant in the Superior Institute of Technologies and the Superior Institute of Computer Sciences and Multimedia of Gabes. In 2006, she joined the Gabes University, as a permanent assistant, where she is currently an assistant professor in the Department of Industrial Computing of the Superior Institute of Computer Sciences and Multimedia of Gabes (ISIMG). Her wide research areas include computer vision and image analysis. Her current research interests are focused on wavelets and wavelet networks and their applications to data classification, image coding. She was the chair of the Workshop on Intelligent Machines: Theories & Applications (17th WIMTA 2010).

Mourad Zaied received the HDR, the Ph.D. degrees in computer engineering and the master degree of science from the National Engineering School of Sfax respectively in 2013, 2008 and in 2003. He obtained the degree of computer engineer from the National Engineering School of Monastir in 1995. Since 1997 he served in several institutes and faculties in University of Gabes as a teaching assistant. He joined in 2007 the National Engineering School of Gabes (ENIG) as where he is currently an associate professor in the Department of Electrical Engineering. His research interests include computer vision and image and video analysis. These research activities are centered around wavelets and wavelet networks and their applications to data classification and approximation, pattern recognition and image, audio and video coding and indexing. He was the chair of the Workshop on Intelligent Machines: Theories & Applications (WIMTA II 2009) and he organized two winter schools on wavelet and its applications (2005) and on Matlab toolkits (2004).