Fuzzy emotion recognition model for video sequences

Oussalah, M.; Wang, S.

DOI:
10.1109/IPTA.2012.6469574

Fuzzy Emotion Recognition Model For Video Sequences

M. Oussalah, S. Wang

University of Birmingham, School of Electronics, Electrical and Computer Engineering
Edgbaston, Birmingham, B15 2TT, UK
M.Oussalah@bham.ac.uk; sxw208@bham.ac.uk

Abstract— **Automatic facial expression recognition from video clips is a challenging task due to computational complexity, limitations of image analysis and subjectivity. This paper advocates a fuzzy based approach for emotion classification. On the other hand, several proposals have been put forward to enhance the pre-processing stage prior to the classification. This includes a combination of a boundary elliptical model for skin detection, adaptive thresholding, principal component analysis and use of cam-shift for face tracking. The performances of the developed system have been evaluated using TFEID and video clips and compared with Bayes' classifier.**

Keywords— **Emotion recognition, Fuzzy inference system, Cam-shift, Face detection.**

I. Introduction

Emotions, such as happiness, sadness and fear, are an important aspect of human intelligence and have been shown to play a significant role on memory, thinking and human decision-making process. This motivates extensive research from various disciplines, e.g., cognitive science, artificial intelligence, to explore various models of emotions as well as their software implementations. Psychologists investigated the role of emotions as a positive component in human cognition and intelligence [1]. On the other hand, increased applications of video-surveillance systems and security related issues has motivated research into automatic emotion recognition systems, especially, it was often observed that automatic or semi-automatic supervision often outperforms the manual supervision due to fatigue effect where a operator(s) would have to monitor screen (s) for 24 hours to in case of missing/ suspicious information. Through years of study, researchers have reached a common ground that the human motion recognition is the key technology for basic human behavior classification, while many researchers have come up with inspiring breakthroughs [2]. But when it comes to areas like studying people's mental health by talking to them or showing some videos and recording their responses, the motion feature is clearly not a significant variable anymore. In this case, the face is the one that carries useful information manifest in the way of facial expressions. Although, its application range is not as wide and not many papers have specifically focused on the construction of this kind of system, there is no doubt that it has its own advantages. Inspired by the psychological models of emotions, Intelligent Agents researchers have begun to recognize the utility of computational models of emotions for improving complex, interactive programs. For instance, interface agents with a model of emotions can provide a better understanding of the user's moods, emotions and preferences and can thus adapt itself to the user's needs [3]. Synthetic characters can use a model of emotion to simulate and express emotional responses, which can effectively enhance their believability [4]. Software agents may use emotions to facilitate the social interactions and communications between groups of agents and help coordination of tasks [5,6]. Emotions can also be used to simulate personality traits in believable agents [7].

On the other hand, the introduction of MPEG standards, which provided an alternative way to analyzing and modelling facial expressions and related emotions [8] through facial animation parameters (FAPs) and facial definition parameter set (FDP), has enhanced the construction of computational model of emotions. are utilized in the framework of MPEG-4 for facial animation purposes. Automatic detection of particular FDPs in a video sequence is an active research area [9], which can be employed within the MPEG-4 standard for analyzing and encoding facial expressions. The strong connection between emotion and facial expression has also been established in Ekman's Facial Feedback Theory" [1], which states that each basic emotion is associated with a unique facial expression and sensory feedback from the expression contributes to the emotional feelings. This motivated later on an intensive research in the field of automatic face detection and facial expression recognition in image vision community. Nevertheless, due to computational complexity, quality of video and images, the task of automatic emotion recognition is still very challenging and open issue in research community.

This paper advocates a fuzzy based approach for emotion recognition. On the other hand, in order to reduce the computational complexity, an image tracking system has been put forward based on camshift algorithm [10]. The approach also makes use of an original method for face detection, which combines an elliptical boundary model for skin color detection [11], adaptive thresholding [12] and principal component analysis (PCA). The feasibility of the proposal has been demonstrated through comparison with standards Bayes' classifier on publicly available database and home-made video clips.

II. SYSTEM OVERVIEW

The video dealt with in this paper is a one-person video and change of the person's position is not as intense as it would be in surveillance field. This simplification is motivated by the desire to focus on the emotion recognition task in timely efficient approach rather than dealing with complex and expensive scene analysis prior to behavior analysis task.

Ultimately, the facial emotion recognition task involves a preprocessing task of face identification. On the other hand, given the importance of the emotion recognition part in this sequel, the time costly processing time of face detection should be minimized. For this purpose, and given the fact that video contains many highly redundant and overlapping frames, a face tracking block were used to track a given face instead of detecting the face at each frame. Besides, considering the fact that face muscles are incapable of moving too quickly, the recourse to face-detection and emotion recognition at each frame is meaningless. Therefore, four or five frames per second would be a reasonable range of proportion of the face-emotion analyzer. Hence, a high hitting rate should be guaranteed while constraining the processing time for each frame in no more than 250 milliseconds. This contributes to achieve a balance between hitting rate and processing time.

From an image processing perspective, given a video sequence as input, one requires to extract those frames which contain a facial image. Next emotion recognition system and then behavior analysis will be run. A generic illustration is given in Figure 1.

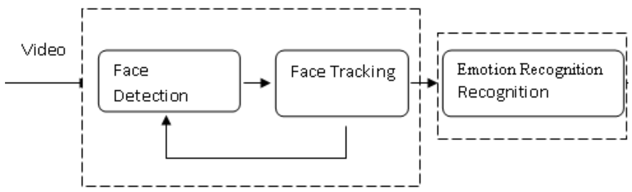

Figure 1. Structure of the Proposed System

III. FACE DETECTION

The face detection task involves two main subtasks. First, face localization method should be used to find out the face region candidates from the image background and other possible objects, to reduce the enormous time cost as well as error rate. Given that skin information is feature invariant and the technique in this field has grown mature, a skin filter based on single Gaussian Model in the same spirit as Feng et al.'s method [13], is constructed to locate face candidates. Next, at each face candidate, a method based on Eigenface feature is used to detect faces.

As mentioned above, it's unwise to detect face for every frame as the position of the face doesn't change much so does the light. Hence tracking the face in the following frames after detecting it can save the time of redundant calculation. For this purpose, a method based on CAM Shift algorithm has been implemented. And if, for some reasons, the tracking fails at one moment, the face detection will be executed again to guarantee the occurrence of face in the underlying image. A block diagram summarization of this approach is highlighted in Figure 2.

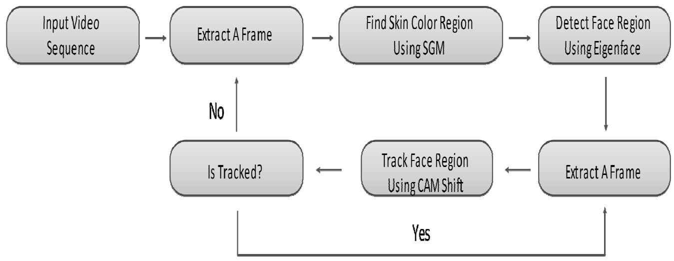

Figure 2. Overview of face detection approach

Single Gaussian Model (SGM) is known to be an efficient way to approximate the distribution pattern of skin color. It is based on the assumption that the color distribution follows a single peak Gaussian distribution, either 2-D or 3-D. The parameters of the Gaussian distribution can be predicted through statistical analysis upon a known skin data set [11, 14]. It is common to use first YCbCr color space when using SGM. This is due to the transformation simplicity and explicit separation of luminance and chrominance components, which makes YCbCr colorspace attractive for skin color modelling. Therefore, one first requires to transform RGB colour into YCbCr space, then uses the Gaussian likelihood model to determine the probability each pixel belongs to skin feature. After color of each region is compared to the skin color to detect potential face regions. The largest detected skin colored region is then processed by the face localization, to verify whether it really contains a face. This task is performed using Eigenface model involving some training database is used to detect the face region. Indeed, Eigenface is aimed at capturing the most significant variances among face images and using this information to encode and compare images of individual face in a holistic or global (opposed to a part-based or feature-based) manner [15]. Principle component analysis (PCA) algorithm is employed to find out those directions with highest variances among data and induces a new space with reduced dimensionality. An image is considered face if the projection distance is larger than some predefined threshold. The Final stage consists in use of Cam shift algorithm in order to track the hue component of an already identified face in the given frame. This avoids the recourse to PCA analysis at each frame to identify face. The basis of the Cam shift tracker [10], which is a variant of mean-shift algorithm that accounts for dynamically changing distribution by readjusting the search window size for the next frame based on the zeroth moment of the current frames distribution. This allows the algorithm to anticipate object movement to quickly track the object in the next scene. Even during quick movements of an object, CAMSHIFT is still able to correctly track.

A. Algorithm and Construction

1) Transform RGB Color Space to YCbCr Color Space

This is achieved through the following standard matrix transformation given the RGB array [R G B]:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 64.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

2) Use Gaussian Algorithm to Obtain Skin Color Likelihood

Given a pixel whose color vector C=[Cb Cr] in YCbCr color space, the likelihood that the pixel belongs to a skin color can be expressed using an elliptical Gaussian joint probability density function:

$$p(C|Skin) = \frac{1}{2\pi|\Sigma|^{1/2}} e^{-\frac{1}{2}(C-M)^T \Sigma^{-1}(C-M)}$$

Where M and $\sum$ are the model parameters (mean and variance-covariance matrix, respectively) of the Gaussian distribution estimated from training data. Higher the probability is, more likely the pixel belongs to skin color; namely, the pixel whose color vector is C corresponds to a skin feature if and only if $p(C\,|\,Skin) \geq \varepsilon$, where ε is a critical threshold that distinguishes skin from non-skin pixel. In order to determine this threshold, a ad-hoc approach has been developed based on Otsu image binarization method. Indeed, OTSU algorithm assumes that the image to be thresholded is composed of two types: foreground and background pixels. It is aimed at finding the optimal threshold value to separate two types of pixels as to achieve maximal inter-class variance and minimal intra-class variance. The overall adapted algorithm is summarized below.

Normalize the probability image to range [0 255] → likelihood image
Compute histogram H [256] of each intensity level
Normalize the H to range [0 1]
FOR (each threshold d in range [0 255])
{
$w_0 \leftarrow \frac{\sum_{i=0}^{d} H(i)}{\sum_{i=0}^{255} H(i)}$ // the ratio of background pixels to total pixels
$w_1 \leftarrow \frac{\sum_{i=d+1}^{255} H(i)}{\sum_{i=0}^{255} H(i)}$ // the ratio of foreground pixels to total pixels
$u_0 \leftarrow \frac{\sum_{i=0}^{d} H(i)i}{\sum_{i=0}^{d} H(i)}$ // the average color of the background
$u_1 \leftarrow \frac{\sum_{i=d+1}^{255} H(i)i}{\sum_{i=d+1}^{255} H(i)}$ // the average color of the foreground
$u \leftarrow u_0 w_0 + u_1 w_1$
Set $G(d) = w_0(u_0 - u)^2 + w_1(u_1 - u)^2 - w_0 w_1 (u_0 - u_1)^2$
}
Choose the threshold D_f corresponding to a value d that maximizes G(d)
Normalize D_f within [0,1]. Set ε as the normalized D_f.

3) Use Eigenface method for face detection and recognition in previously skinned image region

The main idea in eigenface approach is to represent each image into a lower dimensional space constructed from the training images and PCA analysis. Namely, the approach starts with 1-D vector representation Γ_i of image I_i (i=1 to total number of images in training database), and construct $\Theta_i = \Gamma_i - \overline{\Gamma}$ where $\overline{\Gamma}$ stands for mean of Γ_i vectors. Then determine the eigenvectors μ_i of the covariance matrix $C = \sum_i \Theta_i . \Theta_i^T$. The first K "principle axis" corresponding to K largest eigenvalues are then selected (the value of K is chosen so that the sum of the first K eigenvalues provides a large proportion of sum over total eigenvalues.

Now given a new image ζ, or an identified skin-based region of this image, first one constructs 1-D representation ζ_{1D}, then subtract the average face as $\theta = \zeta_{1D} - \overline{\Gamma}$. Next project this image into the elaborated K principle axis as $\theta_p = \sum_{i=1,K} \mu_i^T . \theta . \mu_i$. Finally, the projection distance $d = \left\| \theta - \theta_p \right\|^2$ allows us to classify the underlying image I as face or not depending whether d is larger than some predefined threshold or not.

4) Use of Adaptive Cam Shift algorithm to track identified face frames

This is achieved in a number of steps:

- Get the rectangle region of the face and transform it from RGB color space to HSV (HIS) color space. Then construct a 1-D histogram based on the hue component.
- Transform the color histogram of the original image into color probability distribution image. This process is called the "Back Projection".
- Apply the CAM Shift Algorithm onto the Back Projection Image. This process involves calculating the mass center of the window and approximating it towards the geometric center of the window. The pseudo code is presented below.

mass center ← mass(x,y)
geometric center ← geo(x,y)
FOR each searching window centered at point geo {
//size of the window usually varies from 0.9~1.1 from the original one
$M_{00} \leftarrow \sum_{x,y \in W} I(x,y)$;
$M_{10} \leftarrow \sum_{x,y \in W} I(x,y)\, x$;
$M_{01} \leftarrow \sum_{x,y \in W} I(x,y)\, y$;
$mass.x \leftarrow \frac{M_{10}}{M_{00}}$;
$mass.y \leftarrow \frac{M_{10}}{M_{00}}$;
geo ← mass.copy;
if(||mass – geo||< threshold); //convergence reached
break; }

IV. Emotion Recognition

A. Background

The idea advocated in this paper is based on points-and-angles-based emotion recognition that extends Esau et al.'s

[16] work on extracting Facial Characteristics Points (FCP) and construct a classifier to study the locations or relative positions among them. Such points are those pixels that have unique or significant moving pattern for each emotion. Usually, they are distributed in the areas like eyes, eyebrows, nose, mouth even cheeks. The method employed here makes use of geometric projection. It calculates the summation of the pixel values in horizontal and vertical direction and uses peaks and troughs to determine the location of points of interests, although, one acknowledges its relatively high computational cost. More specifically, after converting the image to gray-level, adaptive thresholding is used for binarization and dividing the image into three main regions: left eye, right eye and nose & mouth. Geometric projection in horizontal and vertical directions is then constructed for each sub region and the FCPs can be extracted by studying the locations of peaks and troughs in the projections. Angles among those FCPs are calculated as properties and main inputs for emotion classification. The choice of angles for classification provides a size invariant classification and saves effort of normalization. Five facial emotions were considered; namely, disgust, happy, neutral, sad and surprise. A Naïve Bayes Classifier (NBC) was employed for this purpose. This will be compared to fuzzy-based approach later on.

B. Algorithm and Construction

1) <u>*Adaptive Thresholding on Gray-level Image*</u>

This is based on Bradley's method [12]. The latter relies on the calculus of the image integral (also known as a summed-area table) and kernel comparison. Especially, with this approach, the face can be divided into three distinct parts: left eye, right eye and nose & mouth as shown in Figure 3.

Figure 3. Result of Adaptive Thresholding

2) <u>*Points Extraction Using Geometric Projection*</u>

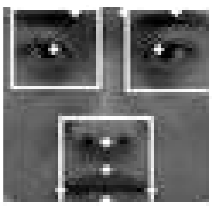

Figure 4. Retrieval of Ten FCPs

Geometric projection is to calculate the sum of the pixel values in a row or column and use this information to represent an image.

After obtaining the three main components in the face, the vertical and horizontal geometric projection is applied to every sub regions. The points of interest can be located by studying the troughs as well as the peaks of the projections. The result is shown in Figure 4.

3) <u>*Angles Retrieval*</u>

After obtaining the FCPs, the angles among them should be calculated. The most commonly used angles are shown in Figure 5 (from A_0 to A_5).

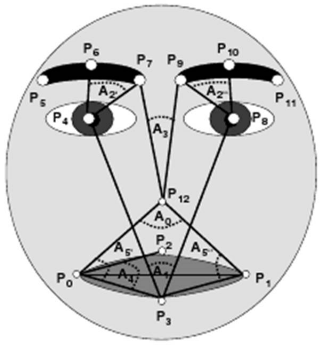

Figure 5. Commonly Used Face Angles

4) <u>*Naïve Bayes Classifier*</u>

For the purpose of a further analysis of the angles retrieved from previous steps, a Naïve Bayes Classifier (NBC) was used due to its simplicity and small training requirement. Although the assumption of the independence among properties is not valid in practical applications, NBC still performs well when the number of attributes is small and the dependence among properties is not too strong which is proven to be applicable in our case. More specifically, angles between fiducial points as depicted in Figure 5 extracted from training set are used to fit Gaussian functions specific for each feature from each emotion class. Then when a new set of angles from a new image is given to the NBC, the probability of each property belonging to each emotion is calculated according to the Gaussian distribution function. The multiplication of those probabilities assigned to each property determines the final likelihood of the image belonging to each emotion. The emotion that has the highest probability would be chosen as the desired one.

V. FUZZY LOGIC APPROACH

The adopted fuzzy like approach is in the same spirit as that in [16]. In this respect, for each emotion class, there is a fuzzy rule system that determines to which degree the actual facial expression belongs to the underlying emotion class based on inputs consisting of geometrical features corresponding to angles (A_0 to A_5). As for conventional fuzzy inference system, the latter involves three main stages (see Fig. 6)

- Fuzzification where for each angle linguistic variable with three linguistic terms small, medium and large is defined, that correspond to the degrees of AU-intensity distinguished by Ekman [1]. See Figure 7.

- Fuzzy inference where the above linguistic variables are evaluated by a set of fuzzy rules that describe the behaviour of the output (emotion) with respect to inputs (linguistic variables).
- Defuzzification where the linguistic (fuzzy) is transformed into crisp value using center of gravity like approach.

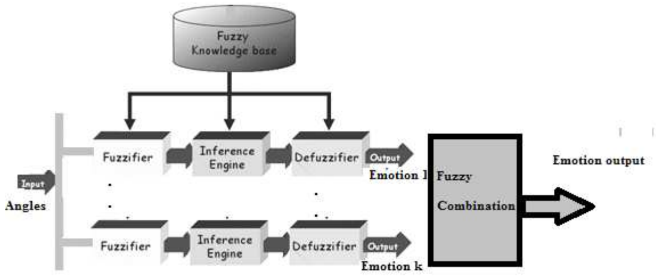

Figure 6. Generic Fuzzy emotion based model

The fuzzification stage, which constructs linguistic quantification of the angles; namely, small, medium and large, is based on the use of trapezdoidal membership functions over the universe of discourse of [0, π] interval whose quantification reflects expert's knowledge about the three above linguistic variables as depicted in Figure 7.

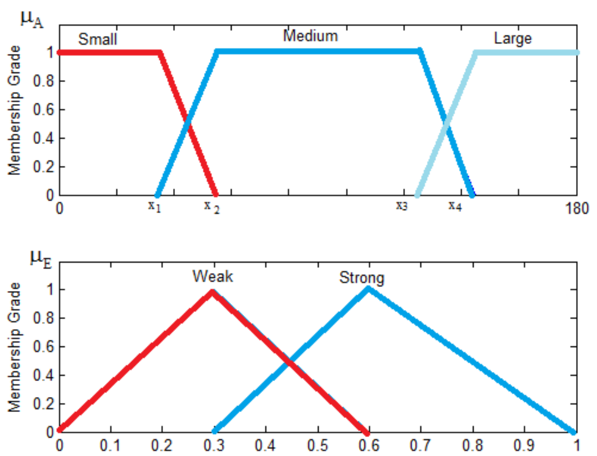

Figure 7. Fuzzification of angles and outputs in the fuzzy system.

The fuzzy rules, which form the basis of fuzzy inference system, are based on intuitive grounds and computational aspect that favor trapezoidal like shapes while the exact values of their support and core as exemplified by parameters x_1, x_2, x_3, x_4 shown in Figure 7 are determined through training stage using part of face database. The premises of each fuzzy rule consist of the linguistic variables A_i (i=0, 5) while its consequent part is constituted of fuzzy emotion value E_j (j=1 to 5, corresponding to disgust, happiness, neutral, sad and surprise, respectively). Two linguistic variables were used to describe the state of each emotion; namely, weak and strong whose membership functions are depicted in Figure 7. Instances of employed fuzzy rules are described in Table 1 (for category strong).

Table 1. Example of Fuzzy rules

	A_0	A_1	A_2	A_3	A_4	A_5
Disgust	Medium	Large	Medium	Medium	Small	Medium
Happy	Large	Medium	Medium	Medium	Small	Large
Sad	Medium	Medium	Medium	Medium	Small	Medium
Neutral	Medium	Medium	Medium	Medium	Medium	Medium
Surpris	Medium	Medium	Medium	Small	Small	Small

For example, if A_0 is large AND A_1 is medium AND A_2 is medium AND A_3 is medium AND A_4 is small AND A_5 is Large, THEN E_{HAPPY} (E_2) is Strong. Therefore, the fuzzy inference system generates a five dimensional vector corresponding to the membership grade, resulting from the defuzzification stage, for each basic emotion. The determination of the dominant emotion state can be obtained using max-like-combination in the case where at least one of the emotion membership grade value is greater than 0.3, indicating the occurrence of at least one strong emotion state, otherwise, the associated motion state will be neutral.

Notice that such approach seems intuitive when there is a single dominant emotion membership grade. However, when membership grades associated to emotion states are very close to each other whose values lie beyond the 0.3 threshold, then the above quantification rule maybe questions in the sense that the choice of the state with maximum membership grade is not effective. However, in practice, the occurrence of such scenarios is minimized thanks to the training phase.

Indeed, the exact shape of the trapezoidal membership functions of input angles; namely, Small, Medium and Large, is tuned to accommodate individual's face. For this purpose, an approach based on least square was developed. More specifically, this boils down to determining the vector

$X = (x_1 \ \ x_2 \ \ x_3 \ \ x_4)$

for each emotion state, such that, the classification rate is maximized. This requires formulating the fuzzy rule activation and quantification, including the defuzzication stage in analytical form that will be cast into nonlinear least square method. In other words, the optimization process is carried out such that the defuzzification value associated to examples belonging to a given emotion state are close to unity.

It should be noted that such optimization approach in tuning the membership grades contrasts with max-min approach employed in [15] for instance. It should be noted that in practice the approach is also limited by the number of examples available in training where only 20 cases were employed for each motion state.

VI. Testing

We used TFEID image database (bml.ym.edu.tw/~download/html/download.htm) where 20 images from each emotion were used for training and the rest of the images for testing. 10 points have been extracted and 7 angles have been used as input to both Naïve Bayes Classifier and fuzzy based system. The results pointed out in Table 2 show the confusion matrix of both approaches where the Bayes' result corresponds to the first element in the table and Fuzzy approach result the second one. For instance, using Bayes' classifier, 14 out of 20 Disgust cases were correctly classified, while 16 out of 20 were correctly classified using fuzzy-based approach.

Table 2. Test Result of Emotion Recognition Approach

	Disg.	*Hap.*	*Neut.*	*Sad.*	*Surp.*	*Total*
Disg.	*14 16*	*0* *0*	*3* *1*	*3* *3*	*0* *0*	*20*
Hap.	*0* *0*	*19* *18*	*1* *2*	*0* *2*	*0* *0*	*20*
Neut.	*4* *2*	*0* *0*	*13* *16*	*3* *2*	*0* *0*	*20*
Sad	*3* *1*	*0* *0*	*3* *3*	*14* *16*	*0* *0*	*20*
Surp.	*1* *0*	*1* *0*	*0* *2*	*0* *0*	*18* *18*	*20*
Tot.	*22* *24*	*20* *20*	*20* *20*	*20* *22*	*18* *18*	*78%* *84%*

Judging from the result, the method can have a good performance on the database we have used. Needless to say, it is based on the assumption that the extractions of FCPs have all succeeded. The features of this method are:

- The processing time is really short, approximately 50ms per frame, which could satisfy the need of real-time processing.
- The point extraction part depends very much on the light condition. Hence, the retrieval of the angles may fail sometimes.
- Even if the varying light condition wouldn't lead to failure when extracting points, the accuracy can't be guaranteed. This ill-located point coordinates can make the result unstable considering the sensitivity of the angles towards FCPs.
- The processing time is really short, approximately 50ms per frame, which could satisfy the need of real-time processing.
- The accuracy of the emotion recognition varies from 65 to 95%. One notices that recognition rate for surprise and happiness achieves higher score of 18/20 (or 90%) and 19/20 (or 95%) while neutral achieves lower score of 13/20 (65%).
- It should be noted that since the emotion states make use of various angles, some of them are more vulnerable to specific geometric features. For instance, the emotions disgust and sadness are more influenced by the upper part of the face (eye brows lifted or knitted and lowered).
- When comparing to Bayes' classifier, it is clear that fuzzy-based approach shows some superiority as testified by the overall classification rate of 84% with respect to 78%. This can be partly explained by the assumption of independence in case of Bayes' classifier as already pointed out, and on the other hand, the flexibility inherent of fuzzy-based system.
- Although the achieved rate is very encouraging and compete with those obtained in the literature, which rarely top 80% rate [2, 15], there is still a room for further improvements.
- This can be achieved by exploring, for instance, the training phase with additional databases, improving the decision making process through refined classification rule that better handles confusing scenarios. The use of Takagi-Sugeno [17] or fuzzy gradual like rules as a substitute to Mamdani-type fuzzy inference may also enhance the classification results.

REFERENCES

[1] Ekman, P. Facial expression and emotion. American Psychologist, 1993, 48 (4): 384–392

[2] J. K. Aggarwal, and Q. Cai, Motion Analysis: A Review. Proceedings of the IEEE Workshop on Non Rigid and Articulated Motion, San Juan, 16 June 1997, IEEE Computer Society Press (90-102)

[3] C. Elliot and J. Brzezinski. Autonomous Agents as Synthetic Characters. *AI Magazine*, American Association for Artificial Intelligence, 1998, 13-30.

[4] J. Bate, The Role of Emotion in Believable Agents. *Communications of the ACM*, 1992, 37(7), 122-125.

[5] M. Domjan. *The principles of learning and behavior* (4th ed.). Boston: Brooks/Cole Publishing Co., 1998

[6] T. Shibata, K. Ohkawa, and K. Tanie. Spontaneous Behavior of Robots for Cooperation - Emotionally Intelligent Robot System-. *Proc. of IEEE Int. Conf. on Robotics and Automation*, Japan, 1998, 2426-2431.

[7] D. Rousseau. Personality in computer characters. In Working Notes of *the AAAI-96 Workshop on AI/ALife*, AAAI Press, Menlo Park, CA, 1996.

[8] M. Tekalp, "Face and 2-D Mesh Animation in MPEG-4," Tutorial Issue On The MPEG-4 Standard, Image Communication Journal, Elsevier, 1999.

[9] Kin-Man Lam and Hong Yan, An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View, IEEE Trans. on PAMI, vol. 20, no. 7, July 1998

[10] J. H. She,, Wang, J.D. and Li P. Face tracking algorithm based on Camshift. Computer Technology and Development, 2008, 18(9): 12-15.

[11] Lee, J.Y. and Yoo, S.I. An elliptical boundary model for skin color detection. In: Proceedings of the International Conference on Image System and Technology, 2002.

[12] Bradley, D. and Roth, G. Adaptive thresholding using the integral image. ACM Journal of Graphics Tools, 2007, 12(2): 13.

[13] Feng, Y.L., Jiang, W.W. and Liu, Z.G. (2010) a novel approach of face detection based on skin color segmentation and PCA. *Science Technology and Engineering,* 2010, 10(10): 2343-2347.

[14] R. L. Hsu, M. Abdel-Mottaleb and A. D. Jain. Face detection in color images. *IEEE Trans. Pattern Analysis and Machine Intelligence 24(5*), 2002, 696-706

[15] Turk, M.A. and Pentland, A.P. (1991) Eigenface for recognition. Journal of Cognitive Neuroscience, 3(1):71-86

[16] Esau, N., Wetzel, E. and Kleinjohann, L. et al. (2007) Real-time facial expression recognition using a fuzzy emotion model, IEEE Proceedings of Fuzzy Systems Conference, London, 23-26 July 2007, IEEE Computer Society Press (1-6)

[17] H. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 1985, 15(1), 116-132.