2012 4th International Conference on Computer Engineering and Technology (ICCET 2012)
IPCSIT vol.40 (2012) © (2012) IACSIT Press, Singapore

Real Time Facial Expression Recognition Based on Multi-feature Integration

Wangyang Ding

Department of Computer Science Tianjin University ,92 Weijin Road, Nankai District, Tianjin 300072, China

wangyang.ding@gmail.com

Abstract. Facial expression recognition is a challenging problem in the fields of pattern recognition and psychology. In this paper, we propose a unified framework to identify and classify facial expressions automatically in real time from a set of images captured by camera. The face area of an image is detected using Adaboost learning algorithm based on Haar-like feature. We extract multiple discriminative texture and geometric features from the mouth and eye areas, which contain most of the facial expression information. After the features are processed using PCA, we train DDAG-SVM to classify the category of the facial expression. Experiments on real face image datasets demonstrate the effectiveness of our method.

Keywords: facial expression recognition, gabor transform, PCA, geometric feature, SVM, multi-feature.

1. Introduction

Facial expression and gesture play an important part of our social interaction. Facial expressions give important clues to understand people's emotions. The psychological research shows that facial expression contains more emotion information than sounds and verbal words. According to Mehrabian's study in 1968 [1] [2] on the three elements account for our liking for the person who puts forward a message concerning their feelings, he gave the equation "total liking = 7% verbal liking + 38% vocal liking + 55% facial liking". As means of human communication, the facial expressions have relatively less variety than the human languages. The connections between facial expressions and emotions have attracted much attention among psychologists. They have put tremendous research efforts to understand the means that the facial expressions are used as body languages for communication, and the general emotions that typical facial expressions indicate across different cultures.

Facial expressions contains profound yet confounding information including recognition (e.g., tired or confused), emotion (e.g., happiness or sadness), reliability, personality (e.g., honesty), special psychological activity and mental disorder. In recent years, facial expression recognition has become a popular research area in artificial intelligence [3][4], especially in artificial psychology theory. Facial expression recognition is a challenging task due to several reasons: 1) The light condition as well as the occlusion of face will affect the recognition performance. Computer uses matrix to process the facial images and it is hard to remove the light effect from the face. 2) Face is a flexible rather than a rigid surface. Expression is given by a sequence of flexible facial muscle motions and related to the facial skin. But the muscle motions and skin textures vary corresponding to people with different ages, genders, races, etc. 3) Traditional image processing methods cannot effectively extract facial expression features.

2. Our contributions

In this paper, we aim to improve the effectiveness and efficiency of facial expression recognition. A novel learning framework is proposed to recognize the expression categories of the facial images in real time. The main steps of our method can be described as follows:

(1) Face detection. We first process the images captured by camera and extract 'Haar-like' features [5]. Trained AdaBoost model is then used to detect the face in the image.

(2) Expression feature extraction. Since most of the expression information are on or near the eyes and mouth, we first use grey integral projection in the vertical direction and some prior knowledge to locate these informative areas. After the extraction of the geometric features, we use Gabor transformation to extract texture feature. The number of features will increase in 10 times as Gabor transformation is used. To meet the real time requirement, we use Principle Component Analysis (PCA [19]) to decrease the feature dimensionality.

(3) Facial Expression classification. Based on the features generated in Step (2), we train DDAG-SVM [15] [22] to classify the face expressions. Cross-validation and grid search methods are combined to turn the parameters of the SVMs.

The rest of the paper is organized as follows: we describe the technical details of our method in Section 3. The experimental design and results are given in Section 4. We finally conclude our work as well as give some interesting future directions in Section 5.

3. Related Work

In 1971, Ekman and Friesen [7] classified six primary emotions which comprises of *happiness, sadness, fear, disgust, surprise* and *anger.* These basic emotion types have been frequently referred to in research literature. Several years later, they proposed Facial Action Coding System (FACS) [8] which use 44 independent Action Units (AU) to describe the facial actions. In the automatic facial expression recognition field, Suwa and Sugie [9] investigated how to analyze facial expressions from an image sequence and presented the results in 1978. In the same year, Waters and Terzopoilos [14] first time generated facial animation based on simplified Ekman-Friesen model. They also presented analysis of the facial expressions.

In recent years, facial expression recognition have been developed starting with the pioneering work of Mase and Pentland [10]. In 1991, they presented an method to use optical flow in 8 different directions to detect the Action Units in FACS, but there was a limitation that no temporal information is described. Yacoob [11] presented a method to extract the optical flows on the mouth, eyes and elbows area and classify the facial expression type using FACS. The method gained very promising recognition performance on 105 face images from 32 persons.

4. Methodology

In this section, we present the details of our learning framework for automatic facial expression recognition in real time. We divide the task into three steps, i.e., *face detection and location*, *facial expression feature extraction* and *facial expression classification*.

4.1. Face Detection and Location

Face recognition research started in 1970s. The research at that time usually adopt template, spatial methods to identify faces in an image [7]. Recent research focus on adopting the data-driven learning algorithms such as statistic model, neural networks, Hidden Markov Model for facial expression recognition. The most widely used algorithm is Adaboost learning algorithm. In this paper, we use the hybrid method proposed by Viola and Jones [5]. The method adopts Haar-like rectangle features and combine them with cascaded Adaboost learning algorithm to detect face.

4.2. Facial Expression Feature Extraction

The ideal facial expression feature extraction method should satisfy:

(1) Extract the facial expression features in a short time.

(2) Eliminate light, noise and camera-related information that is not related to facial expression.

(3) Eliminate redundant facial features and store the features in an compact way.

In this paper, we first transform each image into grey scale and normalize the pixel value to obtain binary balanced grey values after filtering the noise.

The most significant parts account for facial expression are the eyes and the mouth. These two regions contain most informative facial expression features. Therefore, we first locate the regions of eyes and mouth then extract features from these region. This method improves the efficiency of feature extraction by reducing the complexity of pixel analysis and improves the quality of features as well. But this method may lose some important facial expression information by only considering the mouth and eye regions.

4.3. Eye location

In this paper, we use binary grey integral projection method and combine it with prior knowledge to locate the eye and mouth regions. Suppose the grey value is at the location (x, y) of the binary image. The horizontal and vertical projections of the grey integral are $H(x)$ and $V(y)$. $H(x)$ and $V(y)$ can be computed using Eq. (1), (2), where x_{max} and x_{min} are the maximal and minimal values of the horizontal coordinate of the projection, y_{max} and y_{min} are the maximal and minimal values of the vertical coordinate of the projection.

$$H(x) = \frac{1}{y_{\max} - y_{\min}} \sum_{y=y_{\min}}^{y_{\max}} h(x,y) \tag{1}$$

$$V(y) = \frac{1}{x_{\max} - x_{\min}} \sum_{x=x_{\min}}^{x_{\max}} h(x,y) \tag{2}$$

For the sake of simplification, we first reverse the grey image. Suppose the image is 128×128. The vertical position of the left and right eyes can be located using Eq. (3), (4) respectively：

$$Y_e = \arg\max_{y\in[35,65]} \left(\sum_{z=y-2}^{y+2} \sum_{x=20}^{108} h(x,z) \right) \tag{3}$$

$$Y_{er} = \arg\max_{y\in[35,65]} \left(\sum_{z=y-2}^{y+2} \sum_{x=14}^{60} h(x,z) \right) \tag{4}$$

Because both the vertical positions of the elbows and eyes are the trough values. Therefore, suppose Y_{er} is the possible position of an elbow。In next step we determine whether the trough value is the position of an eye. Let Y_{er} move downward 15 pixels. If the following equation holds

$$\sum_{z=Y_{er}-2}^{Y_{er}+2} \sum_{x=14}^{60} h(x,z) > \sum_{z=Y_{er}-17}^{Y_{er}-13} \sum_{x=14}^{60} h(x,z) + 5000 \tag{5}$$

Then Y_{er} in previous step is the eye position, if not, then we need to relocate using

$$Y_{er} = \arg\max_{y\in[45,70]} \left(\sum_{z=y-2}^{y+2} \sum_{x=14}^{60} h(x,z) \right) \tag{6}$$

Based on the center position of the right eye Y_{er}, we then identify the upper and lower boundaries of the right eye. The upper boundary can be identified as in Eq. (7). If there is no Y_{eru} satisfies, then $Y_{eru} = Y_{er} - 25$.

$$Y_{eru} = \min\{y \mid y \in [Y_{er}-15, Y_{er}-30], \sum_{z=y}^{y+1} \sum_{x=14}^{60} h(x,z) < 2000)\} \tag{7}$$

The bottom of the eye can be determined as in (8). If there is no such， then $Y_{erd} = Y_{er} + 20$.

$$Y_{erd} = \min\{y \mid y \in [Y_{er}+5, Y_{er}+20], \sum_{z=y}^{y+1} \sum_{x=14}^{60} h(x,z) < 2000\} \tag{8}$$

Eq (9) calculates the horizontal coordinate X_{er} :

$$X_{er} = \arg\max_{x\in[20,50]} \left(\sum_{z=x}^{x+1} \sum_{y=Y_{er}-15}^{Y_{er}-15} h(z,y) \right) \tag{9}$$

The left boundary of the right eye is given as follows. If there is no such X_{erl}, then $X_{erl} = X_{er} - 25$.

$$X_{erl} = \max\{x \mid x \in [X_{er}-25, X_{er}-10], \sum_{z=x}^{x+1} \sum_{y=Y_{er}-10}^{y=Y_{er}+10} h(z,y) < 1200\} \tag{10}$$

Eq. (11) is used to calculate the right boundary of the right eye. If there is no such X_{err} ,then $X_{err} = X_{er} + 25$.

$$X_{err} = \min\{x \mid x \in [X_{er} + 10, X_{er} + 25], \sum_{z=x}^{x+1} \sum_{y=Y_{er}-10}^{y=Y_{er}+10} h(z, y) < 1200\} \tag{11}$$

Hence we obtain the rectangle area of the right eye, the coordinates of the top left corner and lower right corner are (X_{erl}, Y_{eru}), (X_{err}, Y_{erd}) respectively.

4.4. Mouth Location

The vertical coordinate of the mouth can be located as in the following equation:

$$Y_m = \arg\max_{y \in [84,124]} (\sum_{z=y-1}^{z=y+1} \sum_{x=25}^{103} h(x, z)) \tag{12}$$

The horizontal coordinate can be determined as in the following equation:

$$X_m = \arg\max_{x \in [40,88]} (\sum_{z=x}^{x+1} \sum_{y=Y_m-15}^{Y_m+15} h(z, y)) \tag{13}$$

We can determine the width and the height of the mouth. The width can be calculated by examining the left and right boundary. The x-coordinate of the left boundary can be determined as in (14)，if there is no such X_{ml}, then $X_{ml} = X - 35$.

$$X_{ml} = \max\{x \mid x \in [X_m - 35, X_m - 10], \sum_{z=x}^{x+1} \sum_{y=Y_m-15}^{y=Y_m+15} h(z, y) < 1000\} \tag{14}$$

The right boundary can be determined as in (15). If there is no such X_{mr}, then $X_{mr} = X_m + 35$.

$$X_{mr} = \min\{x \mid x \in [X_m + 10, X_m + 35], \sum_{z=x}^{x+1} \sum_{y=Y_m-15}^{y=Y_m+15} h(z, y) < 1000\} \tag{15}$$

The height of the mouth can be calculated by identify the top and bottom boundary. The x-coordinate of the top boundary can be determined using (16)，if there is no such Y_{mu}, then $Y_{mu} = X_m - 18$

$$Y_{mu} = \max\{y \mid y \in [Y_m - 18, Y_m], \sum_{z=y}^{y+1} \sum_{x=X_m-30}^{x=X_m+30} h(x, z) < 1500\} \tag{16}$$

The bottom boundary can be examined using 17, if there is no such Y_{md}, then $Y_{md} = X_m + 18$.

$$Y_{md} = \min\{y \mid y \in [Y_m, Y_m + 18], \sum_{z=y}^{y+1} \sum_{x=X_m-30}^{x=X_m+30} h(x, z) < 1500\} \tag{17}$$

Based on the above calculations, we can obtain the topleft corner and lower right corner coordinates (X_{ml}, Y_{mu}), (X_{mr}, Y_{md}).

4.5. Feature Extraction from the Eye and Mouth Regions

The geometric features can be extracted very efficiently using automatic location methods [16]. The geometric features related to facial expressions are very few, mainly comprising of the boundaries of the elbows and the distance between the eyes and elbows. The distance between the eyes and elbows can be calculated using the following equation:

$$Y_{mid} = \arg\min_{y \in [Y_{er}, Y_{er}+20]} (\sum_{x=1}^{n} \sum_{j=-1}^{1} h(x + j, y)) \tag{18}$$

After get the coordinate Y_{mid}, we can calculate the distance. In this paper, we use the ten equal intervals along the horizontal direction of the distance between the eyes and elbows.

We search from the center (X_1， Y_{mid}) to examine the binary image upwards and downwards along the vertical direction，if the value doesn't equals to 0, then keep examining. Then we record the number of the consecutive zero-valued pixels. The number is the distance between the eyes and elbows. Then we calculate the numbers of consecutive zero-valued pixels that are started from (X_2, Y_{mid}) … (X_{10}, Y_{mid}) in the same way.

Another geometric feature is the shape of an elbow, which also can be represented using 10 values in the same manner. We can thus get 20 geometric features of the eye regions.

Then we turn to generate geometric features from the mouth region, which mainly contain the boundaries of the upper lip and the lower lip. After we obtain the width and height of the mouth, we can divide the width into 10 equal intervals for both the upper lip and the lower lip and determine the contour lines in each dividing point.

Suppose the upper lip has 10 points, (X_1, Y_{mu}) … (X_{10}, Y_{mu}), along the vertical direction of the 10 points, we can search the pixels downwards that have consecutive values equivalent to 255. We can also search upwards to find the consecutive pixels as well. Moreover, the ratio of the width and the height of the mouth is also an informative feature. We thus obtain 21 features from the mouth region.

4.6. Texture Feature Extraction

The texture features of face contain significant expression information and can represent some subtle facial expression changes. The texture features are informative and the number of them are various.

The Gabor transformation is widely applied to computer vision due to its capability to simulate the vision contour of the mammal's vision neural cells. The Gabor analysis can preserve the spatial and spectral property of an image, which is similar with the analysis of human vision system. Moreover, the Gabor transformation is not sensitive to light condition. Hence, we adopt the Gabor transformation to extract the texture features for facial expression recognition.

Using the Gabor transformation to extract facial expression features is indeed to calculate the convolution of the Gabor filter and facial image. To analyze the image using Gabor filter, we usually adopt a group of Gabor filters with different frequencies in various directions.

Lee [17] and Daugman [18] proposed 2-dimentional Gabor filter group as in the following equation:

$$\psi(x,y,\omega_0,\theta)=\frac{\omega_0}{\sqrt{2\pi}K}e^{-\frac{\omega_0}{8K^2}(4(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2)}\cdot[e^{i(\omega_0 x\cos\theta+\omega_0 y\sin\theta)}-e^{-\frac{K^2}{2}}] \tag{20}$$

where θ represent the direction of the Gabor filter, and ω_0 is the Gabor frequency, the center of the Gabor filter is (0, 0). $K=\sqrt{2\ln 2}(\frac{2^{\phi}+1}{2^{\phi}-1})$ is the ration constant, ϕ is the half bandwidth, $K\approx 2.5$ the frequency bandwidth of 1.5 times frequency interval.

4.7. Facial Expression Classification

The last step of facial expression recognition is to classify the emotion of each facial express ion. We can train classifier based on the resulting PCA features [19].. In this paper, we adopt Support Vector Machine (SVM) [20] [21] due to its robustness and intuitive theory using maximal margin optimization. SVM is an effective statistical learning model in machine learning and data mining. It has been widely applied to biological information, speech recognition, text and image classification, and etc.

However, the basic SVMs only deal with the binary classification tasks, but in real facial expression recognition tasks, there are more than two categories of facial expressions. We hence should use extended SVMs to meet the requirement of multi-class classification.

In this paper, we use DDAG-SVM [22]. DDAG-SVM generates *n(n-1)/2* classifiers, which is the same as the one-versus-one SVM. They also have the same training process. But in the classification process, we need to construct a directed acyclic graph. There are *n(n-1)/2* nodes and n leaf nodes, where each non-leaf node is a classifier, each leaf node is an category of the samples. The classification is processed from the root to a leaf node to get the category of a sample.

5. Experimental Results

In this section, we present the experimental results on real-world facial expression dataset. The experimental environment is as follows:

- CPU: Celeron(R) CPU 2.80GHz
- Memory: 1.0 GB
- Graphics card: ATI RADEN X300 Series
- Operating System: Windows XP 2002 Service 2
- Software: Microsoft Visual Studio 2003. OpenCV 1.0

5.1. System Overview

Real time facial expression recognition system should effectively process the image that is captured by camera and classify the type of the emotion that the facial expression indicates. The facial expression classifier training process in our proposed framework is illustrated in Fig. 5.

In this process, we mainly extract facial expression features and associated the facial expressions to specific emotion type. In the real time recognition process, we capture face image using cameras, then use the classifier learned in the training process to identify the category of the facial expression associated with the specific emotion. The classification results are then output in real time.

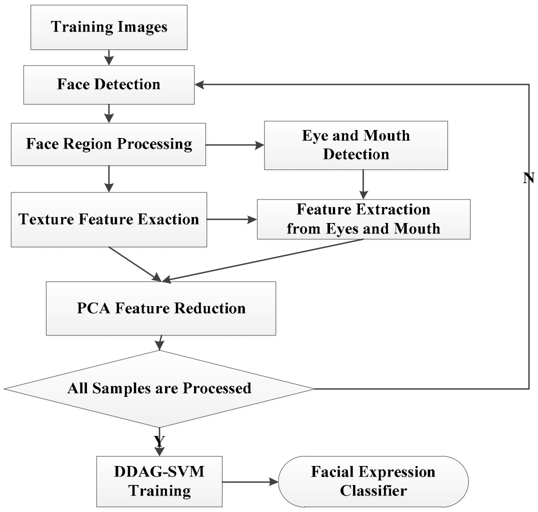

Fig 5: Framework of training facial expression classifier

5.2. Dataset

In the experiments, we consider the popular Japanese ATY dataset (JAFFE)[1]. JAFFE only contains female facial images. There are 213 images of 10 Japanese female models in 7 different expressions. The 7 facial expressions are 6 basic facial expressions (*happiness, sadness, fear, disgust, surprise* and *anger*) plus 1 *neutral*, as shown in Fig. 6. Each female model has 3 or 4 images.

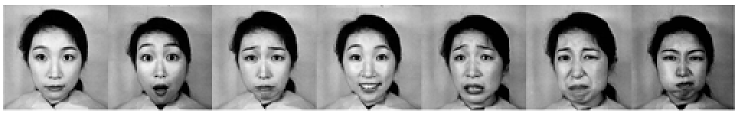

Fig 6: An sample of JAFFE images

5.3. Train and Test

We randomly select 20 images from each facial expression category, hence totally we have 140 images. We train and test the samples using 5-fold cross-validation. In particular, 4 folds are used as training samples, while the left 1 fold is used as testing samples.

We input the training images and use OpenCV method (cvHaarDetectObjects) to detect the faces and output the location and size of the faces. To deal with the problem that the face size is different in each image, we use bilinear interpolation method for normalization. The normalized face size is 128×128. After normalization, we transform the color images into grey images. We then use grey histogram for equalization to improve the contrast ratio. Grey integral projection method is used to locate the eye and mouth regions in the face, where we extract the geometric features. Gabor filter with 4 frequency and 6 directions is used to extract the texture feature, which is reduced using PCA. The resulting texture features and geometric features are integrated for facial expression classification. We train classifiers using DDAG-SVM for recognition.

5.4. Parameter Tuning

The selection of Gabor filter cluster will generate different texture features of the eye and mouth region of the face. There are two parameters in Gabor filter cluster, i.e., the number of frequencies and the number of directions, and we denote as *F* and *D*.

[1] http://kasrl.org/jaffe.html

In SVM, the coefficient *C* determined the ratio of the confidence and the experience risk. It keeps the balance of the generalization ability and recognition accuracy. The parameter γ determines the projection from the original feature space to classifier space. We empirically set $\gamma = 2^{-5}$ and the coefficient C=4. Table 1 shows the facial expression recognition results (accuracy (AC) in percentage and running time in *ms*) for C-SVM which adopts DDAG-SVM classification.

Table1: Parameter Tuning in Gabor filter

D / F	4		5		6	
	AC	Time	AC	Time	AC	Time
3	70. 7	106	72. 1	121	73. 5	131
4	72.1	126	72. 8	142	74. 2	160
5	72. 8	143	74. 2	163	74. 2	196

From Table 1, we can see the DDAG-SCM achieves relative good performance with 4 frequencies and 6 directions . Hence we set the parameters accordingly.

5.5. Results

After turning the parameters, we set *C=4*, $\gamma = 2^{-5}$, $\omega_0 = 2^{-(i+1)}\pi$ *(i=0,...,4)*, $\theta = \frac{j\pi}{6}$ *(j=0,... ,6)*. The recognition results of each category of facial expression (i.e., *happiness(H), disgust(D), sadness(Sa), anger(A), fear(F), surprise(Su)* or *neural(N)*). The results are shown in Table 2. The recognition accuracies are *happiness(85%), disgust(70%), sadness(60%), anger(80%), fear(75%), surprise(90%), neural(70%).* We can see that the model achieve best performance on the category *surprise*. It may be because the discriminative property that faces with surprise exhibit. Moreover, the faces with *sadness* are usually misclassified as that with *fear.* These two categories are actually similar to each other. The overall accuracy *is* 74.28% and the running time is 160ms in total. The results are promising .

Table 2: Results on JAFFE database

	H	D	Sa	A	F	Su	N	AC
H	17	1	0	1	0	0	1	85
D	0	14	2	1	1	0	2	70
Sa	0	1	12	1	4	0	2	60
A	0	1	2	16	1	0	0	80
F	0	1	3	0	13	2	1	75
Su	0	0	0	0	1	18	1	90
N	1	1	2	1	1	0	14	70

6. Conclusions and Future Work

In recent years, there has been increasing research attention on facial expression recognition. In this paper, we aim to build a real time leaning framework for effective facial expression recognition of images. We proposed a method to extract the texture features and geometric features from face image that are discriminative for expression classification. Based on the features, we train DDAG-SVM to classify the category of emotion each facial expression belong to. Experimental results on JAFFE database demonstrate the effectiveness of our proposed learning framework. There are several interesting future directions. One is extending our model to image sequence data such as video. In real world, facial expression is 3D and 3D information can help the recognition task. We plan to focus on the informative feature extraction from 3D face images in the future.

7. References

[1] M.S. Bartlett B. Braathen G.L. Littlewort-Ford J. Hershey J. Fasel T. Mark E. Smith T.J. Sejnowski and J.R. Movellan, "Automatic Analysis of Spontaneous Facial Behavior: A Final Project Report," Technical Report MPLab-TR2001.08, Univ. of California at San Diego, Dec. 2001.

[2] M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski. Measuring facial expressions by computer image analysis. Psychophysiology, 36:25264, 1999.

[3] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," Pattern Recognition, vol. 36, no. 1, pp. 259-275, 2003.

[4] A. Jaimes and N. Sebe. Multimodal Human-Computer Interaction: A Survey. Computer Vision and Image Understanding, 108(1-2):116--134, 2007.

[5] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.

[6] C. Darwin. The Expression of the Emotions in Man and Animals. J. Murray, London, 1872.

[7] Ekman P，Friesen W V. UNMASKING THE FACE：A guide to recognizing emotions from facial clues. Englewood Cliffs, New Jersey： Pretice-Hall Inc, 1975.

[8] P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.

[9] M. Suwa, N. Sugie and KFujimora. A Preliminary Note on Pattern Recognition of Human Emotional Expression. In Proceedings of the 4^{th} International Joint Conference on Pattern Recognition, pages 408-410, 1978.

[10] K. Mase and A. Pentland. Recognition of Facial Expression from Optical Flow. IEICE Transactions, E74(10): 3474-3483, 1991.

[11] Y. Yacoob and M. J. Black. Parameterized modeling and recognition of activities. In Proc. 6th IEEE Int. Conf. on Computer Vision, pages 120-127, Bombay, India, 1998.

[12] I. Essa, A. Pentland. Coding, Analysis, Interpretation, and Recongnition of Facial Expressions. IEEE Transactions On Pattern Analysis and Machine Intelligence, 1997,19(7):757-763.

[13] Philipp Michel, Rana El Kaliouby. Real time facial expression recognition in video using support vector machines. Volume: 2, Publisher: ACM Press, Pages: 258, 2003.

[14] D. Terzopoulos, K. Waters. Analysis and synthesis of facialimage sequences using physical and anatomical model s, IEEE Transactions on Pattern Analysis and Machine Itelligence, 15(6), June, 1993.

[15] Daw-Tung Lina, De-Cheng Pan. Integrating a mixed feature model and multiclass support vector machine for facial expression recognition Integrated Computer aided Engineering 16 (2009) pages 61-74.

[16] D. Cristinacce and T. F. Cootes, "A comparison of shape constrained facial feature detectors," in Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FG '04), 2004.

[17] Lee T. Image representation using 2-D Gabor wavelets [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，1996，18（10）：959-971.

[18] Daugman J G. Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile [J]. Vision Research，1980，20：847-856.

[19] T. Russ, C. Boehnen, and T. Peters. 3D Face Recognition Using 3D Alignment for PCA. In IEEE Conf. on Computer Vision and Pattern Recognition, volume 2, 2006.

[20] R. Fransens, J. De Prins, and L. van Gool, "SVM-based non-parametric discriminant analysis, an application to face detection," in Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV '03), 2003.

[21] Bartlett, M.S., Littlewort, G., Frank, M.G., Lainscsek, C., Fasel, I., and Movellan, J. (2005) 'Recognizing facial expression: machine learning and application to spontaneous behavior', Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conference, Vol. 2.

[22] M. Gonen, A. G. Tanugur, and E. Alpaydin, "Multiclass posterior probability support vector machines," IEEE Trans. Neural Netw., vol. 19, no. 1, pp. 130-139, Jan. 2008.