e-ISSN: 2349-9745
p-ISSN: 2393-8161

Scientific Journal Impact Factor (SJIF): 1.711

International Journal of Modern Trends in Engineering and Research

www.ijmter.com

Research on Emotion Recognition for Facial Expression Images Based on Hidden Markov Model

Jing Mu[1], Fang Chen[2], Changyuan Wang[3]

[1,2,3] *School of Computer Science and Engineering, Xi'an Technological University*

Abstract—This paper introduces emotion recognition for facial expression images using Hidden Markov Model (HMM). Firstly, facial expression images are transformed using discrete cosine transformation and feature is extraction; then HMMs of facial expression images are constructed, and the observation vectors are generated using sub-window, hence, process of emotion recognition for facial expression images is realized. The experiments on JAFFE database are made to recognize the seven emotions of the subjects, the recognition rates are above 80%, so emotion recognition for facial expression images based on HMM is an effective method.

Keywords- Facial expression; Feature extraction; Hidden Markov model; Emotion recognition

I. INTRODUCTION

Facial expression plays a principal role in human interaction and communication since it contains critical and necessary information regarding emotion. Recently, emotion recognition for facial expression images is research hotspot, but the task of automatically recognizing different facial expressions in human-computer environment is challenging.

A variety of system have been developed to perform facial expression recognition, Reference [1] gave a survey on face recognition. The basic method of facial expression recognition is pixel based method. The advantages of pixel intensity values as features are that they are easy to be obtained and they have the same dimensions as the image data. But the disadvantages of pixel values are they tend to be sensitive to image noise as well as image rotations or shifts, and changes in illumination, and they furthermore induce large dimensions on observation vectors. So, some improved method is proposed, such as: geometric feature-based method [2-5]. In Reference [2, 5] Gabor filter are made use of at different scales and orientations, and applied 34 fiducial points for each convolved image to construct the feature vector for representing each facial image. Gabor wavelet coefficients and geometric positions are used to construct the feature vector for each image and applied two-layer perceptron to distinguish seven different facial expressions in [5]. Linear discriminant analysis (LDA) proposed is used to identify seven different facial expressions [6], principal component analysis (PCA) is applied to reduce the dimensionality of feature vectors [7], and independent component analysis (ICA) is proposed [2], and HMM-based face recognition system [8-11]. In the HMM-based face recognition system, each of the significant facial regions such as hair, forehead, eyes, nose and mouth for a frontal face is assigned to a state in a left-to-right one-dimensional continuous HMM.

In the paper, we presents a face recognition system using an HMM approach on the JAFFE (Japanese Female Facial Expression) database [12] because the database is commonly used in measuring the performance of facial expression recognition systems. We use a face recognition system on the JAFFE database to recognize seven main facial expressions: happy, neutral, angry, disgust, fear, sad and surprise. The data compression tool is obtained using the discrete cosine transformation (DCT) for low computation, then an efficient method for extracting the observation vectors using the DCT coefficients is presented. Compared to the classical pixels-based methods, the HMM-based approach offers a more flexible framework for recognition, and can be used more efficiently in facial expression recognition.

The rest of the paper is organized as follows. In Section 2 we describe emotion recognition for facial expression images based on HMM. We make experiments in detail in Section 3. Finally, Section 4 concludes the paper.

II. EMOTION RECOGNITION FOR FACIAL EXPRESSION IMAGES BASED ON HMM

2.1. Description of HMM

The elements of a HMM are described as follows [8]:

1) N , the number of states in the model, if S is the set of states, then we denote the individual states as $S=\{S_1,\ S_2,\cdots,\ S_N\}$, and the state at time t as $q_t, 1\le t\le T$, here, T is the length of the observation sequence.

2) M , the number of distinct observation symbols per state, the set of individual symbols are denoted as $V=\{v_1,\ v_2,\cdots,\ v_M\}$;

3) $\mathbf{A}$, the state transition probability matrix $\mathbf{A}=\{a_{ij}\}$; where $a_{ij}=P[q_{t+1}=S_j \mid q_t=S_i]$, $1\le i,j\le N$, with the constraint $0\le a_{ij}\le 1$, and $\sum_{j=1}^{N} a_{ij}=1, 1\le i\le N$.

4) $\mathbf{B}$, The observation symbol probability distribution in state $j: \mathbf{B}=\{b_j(v_k)\}$, where $b_j(v_k)=P[v_k at\ t|q_t=S_j], 1\le j\le N, 1\le k\le M$.

5) Initial state distribution: $\boldsymbol{\pi}=\{\pi_i\}, \pi_i=P[q_1=S_i], 1\le j\le N$ **.**

A complete specification of a HMM can be as $\Lambda=\{N,M,\mathbf{A},\mathbf{B},\boldsymbol{\pi}\}$. For convenience, using a shorthand notation, a HMM is defined as the triplet $\Lambda=\{\mathbf{A},\mathbf{B},\boldsymbol{\pi}\}$.

2.2 Three basic problems for HMMs

Given appropriate parameterizations of a HMM and an observation sequence as follows: $\mathbf{O}=\{O_1,\ O_2,\cdots,\ O_T\}$, O_t is observation. There are three basic problems described in the following:

Problem 1: Given the observation sequence $\mathbf{O}=\{O_1,\ O_2,\cdots,\ O_T\}$ and a model $\Lambda=\{\mathbf{A},\mathbf{B},\boldsymbol{\pi}\}$, how to $P(\mathbf{O}\mid\Lambda)$, the probability of the observation sequence, given the model, how is it efficiently computed?

Problem 2: Given the observation sequence $\mathbf{O}=\{O_1,\ O_2,\cdots,\ O_T\}$ and model $\Lambda=\{\mathbf{A},\mathbf{B},\boldsymbol{\pi}\}$, how do we choose a corresponding state sequence $\mathbf{Q}=q_1q_2\cdots q_T$ which is optimal in some meaningful sense?

Problem 3: How do we adjust the model parameters $\Lambda=\{\mathbf{A},\mathbf{B},\boldsymbol{\pi}\}$ to maximize $P(\mathbf{O}\mid\Lambda)$?

Problem1 is the evaluation problem, namely given a model and a sequence of observations, we usually use the forward-backward procedure to solve the problem.

Problem 2 is to find the optimal state sequence associated with the given observation sequence. The difficulty lies with the definition of the optimal state sequence. There are several possible optimality criteria. In practice, a formal technique for finding this single best state sequence is used based on dynamic programming methods, and is called the Viterbi algorithm.

Problem 3 is to adjust the model parameters $(\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$ to maximize the probability of the observation sequence given the model, i.e. re-estimation of HMM parameters, the iterative procedure such as the Baum-Welch is usually used to choose $\Lambda = \{\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}\}$ such that $P(\mathbf{O} | \Lambda\}$ is locally maximized.

2.3 Emotion recognition for facial expression images based on HMM

2.3.1 Face expression image HMM

For frontal face images, the significant facial regions (forehead, eyes, nose, mouth, and chin) come in a natural order from top to bottom. Each of these facial regions is a state in a left to right 1D continuous HMM. The state structure of the face model and the non-zero transition probabilities a_{ij} are shown in Figure 1.

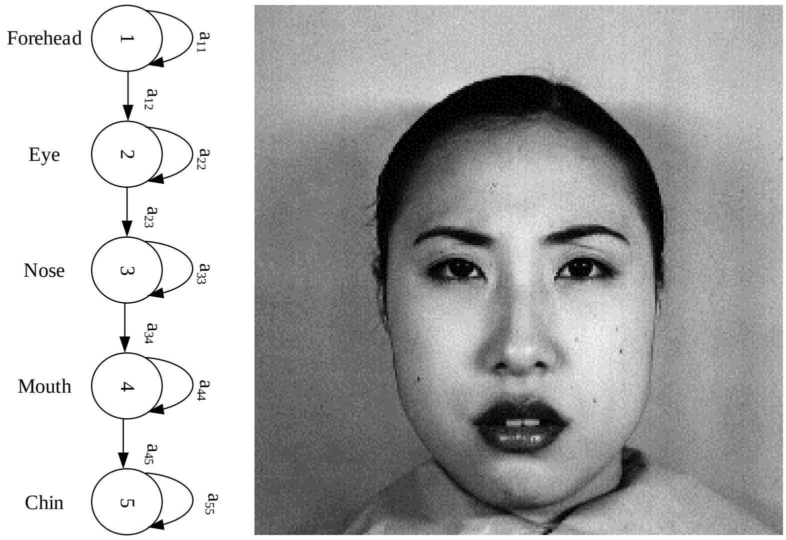

Figure 1. Left to right HMM for face recognition

2.3.2 Feature extraction using discrete cosine transformation

The selection of suitable DCT coefficients from pictures in JAFFE databases was evaluated as a feature extraction method. The DCT transform can be represented as a weighted sum of cosine bases to image. The DCT delivers better energy compression, so the DCT coefficients are always real valued, and the coefficients are nearly uncorrelated. The DCT compression properties and its decorrelation properties make it an attractive technique for the extraction of the observation vectors.

Each face image of width X and height Y is divided into overlapping of height L and width P. The amount of overlap between consecutive blocks is Q and M, as described in Figure 2. We obtain an efficient set of observation vectors using DCT coefficients extracted from each block.

Using DCT transformation, we can obtain a DCT coefficient matrix for image block. These coefficients represent the energy contribution by different frequencies. The first coefficient $C(0,0)$ represents the "DC" component. The rest of the coefficients represent the different "AC" components, as contributed by each of the frequencies present. The compression property of the

DCT allows a block of pixels to be represented by just a few DCT coefficients, and still obtain more information than would be present. In order to extract the coefficients which contain the most data about the block of data transformed, the DCT coefficient matrix needs to be scanned in a zig-zag pattern. This is because the contributing frequencies are arranged from low to high as indicated by the zig-zag pattern.

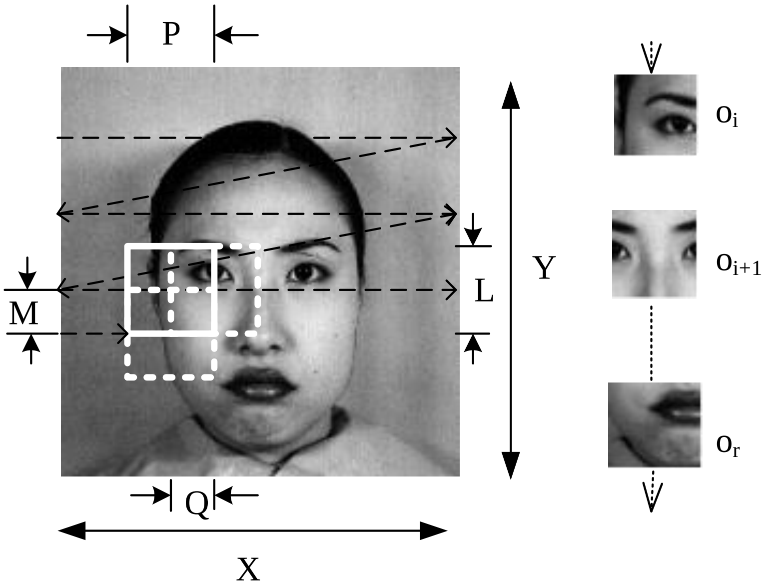

Figure 2. Face image parameterization and block extraction

In this paper, we use a sliding window of 8×8 pixels to scan over a picture with the standard overlap of 50% in both the horizontal and vertical directions. For each window of 8×8 pixels, a DCT coefficient matrix of the same size was obtained. This means that an image of Y rows and X columns there are

$$N_D = (2\frac{X}{N} - 1)(2\frac{Y}{N} - 1) \tag{1}$$

Number of 8×8 DCT coefficient blocks (with N=8 being the size of the window). These DCT coefficient blocks are the reduced by keeping their first 15 coefficients [9] by following the zig-zag pattern. Thus every 64 values are reduced to $L = 15$ values and single observation used to represent the data of block (a, b) is now the vector:

$$\mathbf{X} = [c_0^{(a,b)} \quad c_1^{(a,b)} \quad \cdots \quad c_{L-1}^{(a,b)}] \tag{2}$$

A complete observation sequence is obtained consisting of N_D these vectors. According the observation sequence obtained, we trained the HMMs; then we recognize the test images.

III. EXPERIMENTS

The image database we use in our experiments is the JAFFE [12]. The database contains 213 grayscale facial expression images of ten young Japanese females. The ten Japanese females are numbered as KA, KL, KM, KR, MK, NA, NM, TM, UY, and YM. There are seven different facial expressions, such as neutral, happy, angry, disgust, fear, sad and surprise. Each female has two to four examples for each expression. Each image is of size 256×256. Many References such as [2, 5-7] used this database as the benchmark.

Now, we use HMM based method to recognize the emotion for facial expression images in the JAFFE database. The experimental results are showed in Table 1. Figure 3 presents some of the recognition results. The crossed images present incorrect classifications, while the rest of images are examples of correct classification.

Table.1 shows that the recognition rates of the method are above 90% to the emotion of facial images for KA, KL, KM, MK, NM, TM, the reason is that those females have the exaggerated facial expressions, and exaggerated facial expressions have much information, the difference of different facial expression images is quite obvious, so the recognition rate of the method is high; the recognition rate of the method to KR, NA, UY is above 80%, but to YM is only 63.6%, because the facial expressions of YM has little change and have less the effective information, such as "fear" is recognized to "sad", and "Happy" is to "sad". From Table 1, we see that the total recognition rate is 87.3%, and this demonstrates that the HMM base method is an effective algorithm.

TABLE 1. Emotion recognition rate of JAFFE database

Personal No	Recognition Rate
KA	91.3%
KL	95.5%
KM	90.9%
KR	85%
MK	95.2%
NA	85.7%
NM	100%
TM	95.2%
UY	80.1%
YM	63.6%
Total recognition rate	87.3%

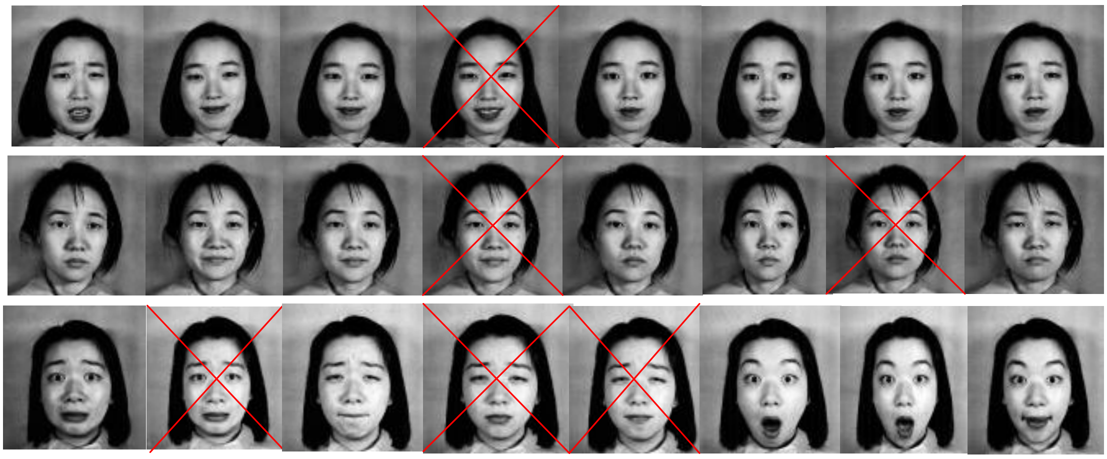

Figure 3. Wrong classifications on facial expression images from JAFFE database using HMM based method

IV. CONCLUSIONS

We investigate DCT based feature representation, which uses an efficient set of observation vectors based on the extraction of the DCT coefficients, then use the HMM based expression classification schemes to recognize seven emotions for facial images on the JAFFE database.

Experimental results show that the recognition rate of the system is 87.3%, the proposed system using HMM based method has better performance. So the HMM modeling of human faces appears to be a promising method for face recognition under a wider range of facial expression, such as aeronautics and astronautics, medical treatment of patients, human-computer interaction and emotional robot, and has the potential value.

ACKNOWLEDGMENT

The authors would like to thank the support of the Key Foundation of Xi'an Technological University under Grant XAGDXJJ1219 and Special fund of Educational Department of Shaanxi Province under Grant 14JK1347.

REFERENCES

[1]. R.Chellappa, C.Wilson, and S.sirohey. Human and machine recognition of faces: a survey. Proceedings of IEEE.vol.83, no.5, pp. 705-740, 1995.

[2]. I.Bucia, C.Kotropoulos and I.Pitas, ICA and Gabor representation for facial expression recognition, Proc. IEEE Int.Conf.Image Processing, vol.2, pp.855-858, 2003.

[3]. Q.Liu. Research on facial expression recognition based on Gabor. Suzhou University, 2009.

[4]. H.Lu, Research on facial feature extraction and recognition in automatic facial expression analysis system. Southeast University, 2006.

[5]. Z. Zhang, M. Lyons, M. Schuster and S. Akamatsu, Comparison between geometry based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron, Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, Nara, pp.454-459, 1998.

[6]. M.Lyons, J.Budynek and S.Akamatsu, Automatic classification of single facial image. IEEE Trans. Patt.Anal. Mach.Intell. vol.21,no.12, pp.1357-1362,1999.

[7]. S. Dubuisson, F. Davoine and M. Masson, A solution for facial expression representation and recognition, Sign. Process: Imag. Commun. vol.17,no.9, pp. 657-673,2002.

[8]. L.Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, vol.77, no.2, pp.257-286, 1989.

[9]. J. Ballot. Face Recognition Using hidden Markov models, Master, University of Stellenbosch, 2005.

[10]. C. Sanderson. Face processing &frontal face verification. IDIAP-RR 20, IDIAP, 2003.

[11]. F. Samaria. Face recognition using hidden Markov models, University of Cambridge, 1994.

[12]. M. Lyons, M. Kamachi and J.Gyoba. Japanese Female Facial Expressions (JAFFE), Database of digital images, 1997.

