Facial Expression Recognition using Local Binary Patterns and Kullback Leibler Divergence

AnushaVupputuri, SukadevMeher

***Abstract*— Facial Expressions play major role in interpersonal communication and imparting intelligence to computer for identifying facial expressions is a crucial task. In this paper we present an efficient preprocessing algorithm combined with feature extraction using Local Binary Patterns (LBP) followed by classification using Kullback Leibler (KL) divergence. Firstly Viola Jones algorithm is used to detect pair of eyes using which effective part of face is obtained which is further processed to eliminate illumination effect. LBP operator is then applied on the preprocessed image to extract local features represented by histogram. Template histograms for seven basic expressions using training images are formed which are compared with the test histogram distribution using an efficient KL divergence for dissimilarity measure. This algorithm is implemented on JAFFE database resulting in a high classification accuracy of 95.24%.**

***Index Terms*— Facial expression, face model, LBP, KL divergence.**

I. Introduction

Facial expression is an immediate and effective part of communication among humans in conveying their internal emotions and intentions. Recognizing facial expressions is an important and efficient task as it finds wide applications in areas of Human Computer Interaction (HCI), driver state monitoring and aiding autistic children. As the facial expressions are combinations of many emotions we categorize them into seven major categories of anger, disgust, fear, happy, sad, surprise and neutral. These expressions are differentiated by very fine variations in muscular movements and hence local feature extraction to represent expressions is a critical task. Prior to this, in real time images or video sequences the face occupies only a small area and extraction of this face part followed by other preprocessing steps plays a major role. Previously researchers extracted facial area based on skin color as the parameter [11] which proved inefficient for gray scale images, hence we adopted Haar like features followed by Adaboost algorithm which is a cascade of simple features to differentiate face class from non face class [4] .To eliminate illumination effect we followed efficient preprocessing technique [6]. Gabor wavelet based feature extraction [5][9] is effective in face representation but size of feature vector so obtained being large hinders effective computation.

Local binary patterns [1] have proved efficient enough in texture analysis and have been extended to facial feature representation for face recognition [3].Their ease in computational complexity and robustness to illumination variations proves their significance for face representation in form of histogram. Support vector machine (SVM) [10] though being a binary classifier was extended to multiclass classification but its complex training system makes it unsuitable for high speed applications. Minimum distance classifiers like Euclidean distance are usually implemented for comparison of the histograms. Here we took the advantage of Kullback Leibler (KL) divergence for histogram comparison.

This paper proposes a facial expression recognition method by applying LBP operator [2] on preprocessed images to extract features represented as a histogram. Template histograms for each expression class thus obtained are compared with test image feature histogram using KL divergence which gives minimum divergence with target class. This method is trained and tested on JAFFE database. The remaining paper is organized as follows. Preprocessing algorithm and LBP feature extraction are explained in sections II and III respectively. Section IV describes the proposed method. Experimental results comparing different distance classifiers on JAFFE database conclude the paper.

II. Face Representation using LBP

Local Binary patterns have proved efficient in texture analysis and are extended to facial feature representation for recognition. The original LBP operator was introduced by Ojala et al. [1] where the center pixel of a 3×3 window is used for thresholding surrounding pixels resulting in an 8 bit binary number. This is illustrated in Fig.1.Each LBP is considered as a micro texton as it codifies micro patterns like curves, edges, spots which are essential features [8] describing a face. Generated LBP image in Fig. 2 portrays the efficiency of LBP in representing regions essential for expression coding.

An extension to LBP operator called uniform LBP accounts to about LBP operator is said to be uniform if there are at most two bitwise transitions from 1 to 0 or vice versa, the binary string being considered circular. For example,

Anusha Vupputuri is with the Department of ECE, National Institute of Technology, Rourkela (Odisha), India (e-mail: anu73vuppu@gmaiI.com).

Sukadev Meher is with the Department of ECE, National Institute of Technology, Rourkela (Odisha), India (e-mail: sukadevmeher@gmaiI.com).

00000000(0 transitions); 01000000, 11111110; (2 transitions) are uniform binary patterns. Accumulating the patterns which have more than 2 transitions into a single bin yields an LBP operator, denoted by $LBP_{P,R}^{u2}$ where $u2$ stands for uniform pattern , P is the number of sampling points equally spaced on a circle of radius R. In an image I all uniform pattern following pixels are labelled individually and those that are non uniform are given a single label. For example, the number of labels for a neighborhood of 8 pixels is 256 for the standard LBP but 59 for $LBP_{P,R}^{u2}$.Thus we obtain labeled image $f_l(x, y)$.The histogram of the labeled image $f_l(x, y)$ is given by equation (1)

$$H_i = \sum_{x,y} I(f_l(x, y) = i),\ i = 0,.........., n-1 \qquad (1)$$

where $I(A) = \begin{cases} 1\ldots if\_a\_is\_true \\ 0\ldots if\_a\_is\_false \end{cases}$

and n is the number of different labels generated by the LBP operator.

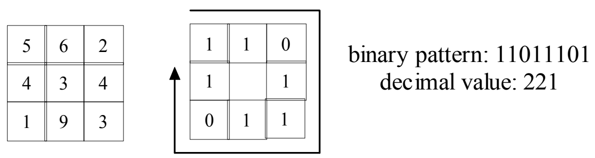

Fig.1. Center pixel thresholding for 3×3 window

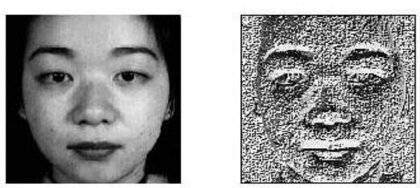
Fig.2. Example image and its generated LBP image

Facial expressions vary with very slight muscular movement and LBP concentrates on these movements using labeled micro patterns which are distributed throughout the facial region. Moreover facial expression variations are typically characterized by areas like corners of mouth, inner corner of eyes and chin, and other sub regions being slightly important. For comparing two images we also need to consider the spatial information of the micro patterns. In order to deal with this aspect the image is subdivided into regions $R_1, R_2, ..., R_m$ which are allotted incremental weights as shown in Fig.3.The histogram of this subdivided image including spatial information is given by equation (2) and the generation of LBP feature histogram from the combination of sub regions is shown in Fig. 4.

$$H_{i,j} = \sum_{x,y} I(f_l(x, y) = i) I((x, y) \in R_j) \qquad (2)$$

where $i = 0,........, n-1$ and $j = 0,........, m-1$

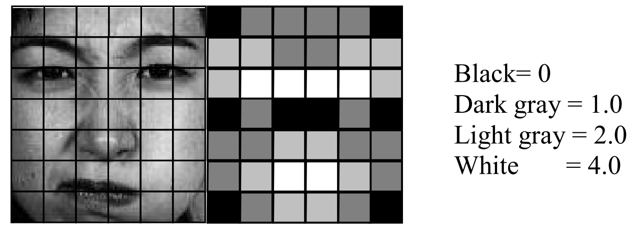

Fig.3. Incremental weight allotment to sub regions

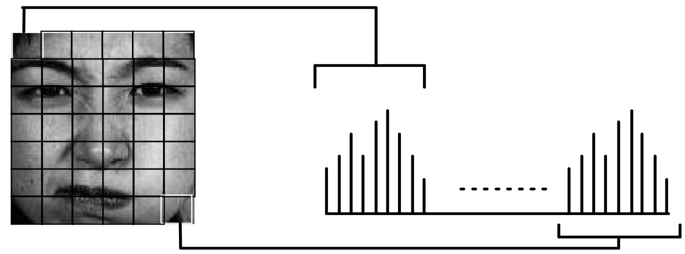
Fig.4. LBP feature histogram

Each image is subdivided into 7×6 regions and each sub region has uniform LBP range of labeled pixels from 1 to 59 which are the corresponding 59 bins for subpart of feature histogram as shown in Fig. 4.Thus $LBP_{(8,1)}^{u2}$ is implemented to obtain feature histogram for comparison.

III. Classification using KL Divergence

Mathematical representation of LBP facial features is available in form of histogram which is best compared using distance classifier. Support Vector Machine (SVM) based classification involves complex training and is not implementation worthy. ANN classifier requires repeated experimentation due to its lack of consistency; different distance classifiers or similarity measures which can possibly yield better results for histogram comparison are as follows:

- *Kullback Leibler (KL)* divergence:

  It is a non symmetric measure of estimating the dissimilarity between two probability distributions here histograms. It gives the information lost while mapping one distribution to other, therefore greater the loss higher is the divergence or dissimilarity between the distributions. KL divergence [7] is calculated using equation (3) shown below.

$$KL(S, M) = \sum_{i,j} w_j S_i \log \frac{S_i}{M_i} \qquad (3)$$

where S, M are the two distributions under comparison for dissimilarity.

- *Weighted Histogram intersection (HI)* which is maximum distance classifier given by equation (4)

$$D(S,M)=\sum_{i,j} w_j(S_i,M_i) \tag{4}$$

- *Euclidean Distance (ED)* which is minimum distance classifier given by equation (5)

$$D(S,M)=\sqrt{\sum_{i,j} w_j(S_i-M_i)^2} \tag{5}$$

- *Chi Square (CS)* statistic dissimilarity measure classifies with minimum dissimilarity value given by equation (6)

$$\chi_w^2(S,M)=\sum_{i,j} w_j \frac{(S_{i,j}-M_{i,j})^2}{S_{i,j}+M_{i,j}} \tag{6}$$

- *Log Statistic* (LS) measure is maximum distance classifier given by equation (7)

$$D(S,M)=-\sum_{i,j} w_j S_i \log M_i \tag{7}$$

IV. PROPOSED METHOD

Proposed method for expression classification is shown in Fig. 7. and these steps are followed by their description and implementation

A. Preprocessing

Preprocessing is an essential and effective step to enhance real time images for recognition. Face occupies a considerably small part of image or video frame. To extract this part efficiently we utilized Viola Jones algorithm which applies the concept of integral image for fast feature evaluation. Adaboost algorithm selects essential reminiscent of Haar features which form weak classifiers. Cascading these weak classifiers to a strong one differentiates face class from background effectively. Firstly pair of eyes is extracted and effective face area is cropped by a face model [12] as shown in Fig. 5. Eye feature points are detected and if distance between them is d then face area is $2.2d\times1.8d$. To eliminate illumination effect, local shadowing and highlights Tan [6] proposed simple steps which are near to mammalian visual cortex. They include Gamma correction I^{γ} where $\gamma=0.2$ on the image I to enhance local dynamic range of image in dark regions and compressing in lighter regions. Shading is low frequency information which can be removed using Difference of Gaussian (DoG) filtering followed by contrast equalization to globally rescale the images in order to improve overall contrast using equation (8) and equation (9). Tan hyperbolic function is applied further to limit the range of intensities. Illustration of these steps is shown in Fig. 6.

$$I(x,y)=\frac{I(x,y)}{(mean(|I(x',y')|^{\alpha}))^{\frac{1}{\alpha}}} \tag{8}$$

$$I(x,y)=\frac{I(x,y)}{(mean(\min(\tau,|I(x',y')|)^{\alpha}))^{\frac{1}{\alpha}}} \tag{9}$$

Fig.5.Cropped face from JAFFE database using Face model and Viola Jones algorithm

(a) (b) (c) (d)

Fig.6. Resultant images after (a) gamma correction, (b) DoG filtering, (c) contrast equalization and (d) tan hyperbolic limiting (from left to right)

B. LBP feature representation

Preprocessed Training data set images are grouped into individual emotion classes and uniform LBP operator using 3×3 window is applied on them to represent face using LBP feature histogram. Feature histograms of images corresponding to a single class are averaged to form template histogram of that particular emotion class.Thus 7 template histograms corresponding to each emotion class are obtained which are further used for comparison. Similar procedure is followed for test images.

C. Minimum distance classification

Kullback Leibler (KL) divergence is a minimum distance classifier efficient in comparing two probability distribution hence we use it here for histogram comparison. Test image histogram is compared with template histogram of each class and the one giving minimum divergence is determined as detected emotion class.

D. Database description

This paper uses JAFFE database for training and testing the algorithm. JAFFE is an acronym for Japanese Female Facial Expression. JAFFE database provides 213 images posed by 10 Japanese females showing all the seven basic emotions (Anger, Disgust, Fear, Happy, Sad, Surprise and Neutral). The database was planned and assembled by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba [13].

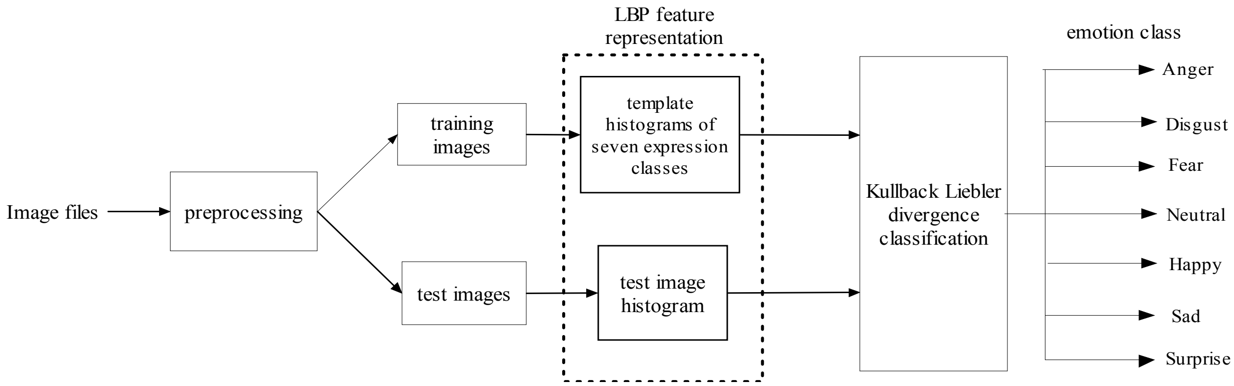

Fig.7. Proposed method for facial expression recognition

V. EXPERIMENTAL RESULTS AND COMPARISONS

The proposed algorithm was trained and tested on the 210 images of JAFFE database. Images of different subjects belonging to a single emotion class are grouped together to form the training database. One subject out of ten is chosen to form the test images. The training was performed in three rounds, choosing varied test subjects each time. Individual emotion class accuracy and overall classification accuracy with and without preprocessing are compared and is shown in following Table I.

TABLE I: CLASSIFICATION ACCURACY OF PROPOSED METHOD WITH AND WITHOUT PREPROCESSING

Classification method	**Emotion class**						
	AN	***DG***	***FR***	***HA***	***NT***	***SD***	***SU***
without preproc.	66.7	66.7	66.7	**100**	**100**	33.3	**100**
With preproc.	**100**	**100**	**100**	**100**	**100**	**66.7**	**100**

The abbreviations AN, DG, FR, HA, NT, SD, SU represent anger, disgust, fear, neutral, sad and surprise respectively.

KL divergence used for classification in the proposed method is compared with other distance and dissimilarity measures in terms of overall classification accuracy. KL divergence yielded a highest accuracy of 95.24% which is considered to be greater than other distance based classification methods mentioned in Section III. KL distance measures for weighted person independent classification is shown in Table II. Images from test set are randomly chosen and were observed that image belonging to a particular emotion class gave minimum KL distance against that particular class.

TABLE II: KL DIVERGENCE MEASURE FOR SEVEN EMOTION CLASSES

I/P Emotion class	**KL divergence values**						
	AN	***DG***	***FR***	***HA***	***NT***	***SD***	***SU***
AN	**1.2173**	1.3424	1.4496	1.4423	1.4912	1.3746	1.6481
DG	1.2565	**1.2438**	1.2947	1.3423	1.4488	1.2788	1.5139
FR	1.4094	1.3125	**1.2358**	1.387	1.3424	1.2366	1.3638
HA	1.8282	1.74	1.8477	**1.2755**	1.9625	1.7982	2.0466
NT	1.4629	1.5394	1.4506	1.4836	**1.3105**	1.3805	1.4499
SD	1.6229	1.4161	1.3870	1.5304	1.5673	**1.3272**	1.6165
SU	1.6341	1.7622	1.5730	1.8652	1.5442	1.5688	**1.2594**

Further KL divergence is comparison with other distance classifiers is shown in Table III. From this we can infer that our proposed method which is a proper combination of preprocessing technique, LBP feature representation with KL divergence comparison is accurate and efficient for expression classification.

TABLE III: COMPARISON OF VARIOUS DISTANCE CLASSIFIERS

Emotion class	**Distance classifier**				
	HI	***ED***	***CS***	***LS***	***KL***
Anger	66.7	66.7	66.7	100	100
Disgust	66.7	66.7	33.3	100	100
Fear	33.3	66.7	33.3	66.7	100
Happy	100	100	100	100	100
Neutral	66.7	66.7	66.7	66.7	100
Sad	33.3	33.3	33.3	66.7	66.7
Surprise	66.7	100	100	100	100
Overall accuracy	62	71.44	62	85.72	**95.24**

VI. CONCLUSION

An efficient facial expression recognition method is proposed in this paper which uses robust Local Binary Patterns for facial feature extraction and representation along with Kullback Leibler divergence for classification. Preprocessing improves the classification accuracy of KL divergence measures. Experimental results show that preprocessing combined with KL divergence have less confusion for the correct emotion class recognition. They also show that the proposed method gives a classification accuracy of 95.24% which is better than other distance based classification methods whose accuracy is ranging between 62%-85.72% depicting classification improvement of 9.99% to maximum of 34.9% over existing HI and LS classifiers. Although proposed method proves in terms of accuracy, there still lies confusion with classifying sad and fear classes because single emotion is a combination of many intensions whose recognition becomes simple accompanied with a combination of speech and gesture recognition. The proposed method can be extended to dynamic facial expression recognition from video sequences and implemented for real time situations as both preprocessing and LBP are robust to illumination variations. Our proposed method combining LBP with KL divergence with high classification accuracy delineates its contribution in the field of facial expression classification.

ACKNOWLEDGMENT

The authors would like to thank Dr.Michael Lyons, Miyuki Kamachi, and Jiro Gyoba for providing Japanese Female Facial Expression (JAFFE) database.

REFERENCES

[1] Rein-Lien Hsu; Abdel-Mottaleb, M.; Jain, A.K., "Face detection in color images," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.24, no.5, pp.696,706, May 2002

[2] Paul Viola, Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features" , *IEEE Conference in Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001.

[3] Tan X., and Triggs B., "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions", Proceedings of the 2007 IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'07), Rio de Janeiro, Brazil,pp.168-182, Oct 2007.

[4] Michael, Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba "Coding Facial Expressions with Gabor Wavelets" *Proceedings, Third IEEE InternationalConference on Automatic Face and Gesture Recognition, ,Nara Japan, IEEE Computer Society*, 200-205, April 14-16,1999.

[5] Wenchao Zhang; Shiguang Shan; Wen Gao; Xilin Chen; Hongming Zhang, "Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition,"*Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* , vol.1, no., pp.786,791 Vol. 1, 17-21 Oct. 2005

[6] T Ojala, M Pietikanen and T Maenpaa, "Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, 24(7), pp.971-987.

[7] T. Ahonen, A. Hadid and M. Pietikainen, "Face Recognition with Local Binary Patterns", The 8th European Conference on Computer Vision, 2004, pp.469-481.

[8] Ching-Chih Tsai, You-Zhu Chen and Ching-Wen Liao, "Interactive Emotion Recognition Using Support Vector Machine for Human-Robot Interaction", *Proc. IEEE International Conference on Systems, Man, and Cybernetics,* pp. 407-412, 2009.

[9] Shan C., Gong S., P. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," *Image and Vision Computing,* vol. 27(6), pp. 803–816, 2009.

[10] L. Zhang and D. Tjondronegoro, "Facial Expression Recognition Using Facial Movement Features" *,IEEE Transactions on Affective Computing,*vol.2(4), pp.219–229, 2011

[11] Kullback, Solomon, and Richard A. Leibler. "On information and sufficiency."The annals of mathematical statistics (1951): 79-86.

[12] QUAN-YOU ZHAO, BAO-CHANG PAN and JIAN-JIAPAN, "Facial expression recognition based on fusion of Gabor and LBP features", Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, Hong Kong, 30 31 Aug. 2008,vol 1,pp.362-367.

[13] Lyons, Michael J., et al. "The Japanese female facial expression (JAFFE) database." *1998-12-20)[2010-02-27]. http://www. ntis. atr. co. jp/N mlyons/jaffe, html* (1998).