Facial Expression Recognition for Color Images Using Local Binary Pattern and Log Gabor Transform

S.Arivazhagan, R.Newlin Shebiah, V.A.M. Fathima Benazir and P.Monica Selastine

Abstract*:* Facial expressions convey non-verbal cues, which play an important role in interpersonal relations. It is also used in behavioral science and clinical practice. Automatic recognition of facial expressions is an important component of natural human-machine interfaces. Reliable expression recognition by machine is still a challenge. The main objective of this paper is to investigate two different methods of feature extraction for person-independent facial expression recognition from images. The logarithmic Gabor filters and the local binary pattern operator (LBP) were used for feature extraction. The features were classified using support vector machine (SVM) classifier. The proposed system is designed and tested with Positioning Global Oversight database which consists of 24 different object performing two different expressions are present. Thus a total of 55 images are present. Experiments carried out on Positioning Global Oversight database showed comparable performance between Log-Gabor filters and LBP operator, with a classification accuracy of around 98.96% and 96.87% respectively.

Keyword: **LBP operator, Facial Expression Recognition, support vector machine classifier, Log Gabor transforms.**

I.INTRODUCTION

Facial expression recognition plays an important role in variety of applications like automated tools for behavioral research, bimodal speech processing, video conference, airport security and access control, building (embassy) surveillance/monitoring, human computer intelligent interaction and perceptual interfaces, etc. But automatic FER system needs to solve the following problems such as detection and location of faces in a cluttered scene, facial feature extraction, and facial expression classification [1]. one of the most active and visible research topics in computer vision, pattern recognition and biometrics, face recognition has been extensively studied in the past two decades [2], yet it is still a challenging problem in practice due to uncontrolled environments, occlusions and variations in pose, illumination, expression and aging, etc. Chen,et.al, [2]used a pre-filtering approach to detect candidate

Regions and then used a hybrid set of features consisting of Haar-like and Gabor feature to train various classifiers for faces in multiple poses.

S.Arivazhagan, R.Newlin Shebiah, V.A.M. Fathima Benazir and P.Monica Selastine are with Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi - 626 005, Email: fathima122benz@yahoo.com contactatmoni@gmail.com

This causes a extra computational load to the system.

A face detection technique was presented by M.H.Yang et.al [3], which discusses a different face detection technique based on neural networks, hidden markov based models, template matching method, knowledge based methods, but none of them works in real time applications and they are computationally too expensive.

Cohen et.al,[4] compared Bayes classifiers and Gaussian Tree-Augmented-Naive (TAN). Among the two, Bayes classifiers performed the best result. He also proposed a multi-level HMM classifier, which allows not only to perform expression classification on a video segment, but also to automatically segment a long video sequence to the different expressions segments without resorting to heuristic methods of segmentation. But HMMs do not deal dependencies.

M.Pantic and Rothkrantz [5] presented a detail survey on facial expression recognition of static facial images. Here facial images were extracted and classified using Template based methods, Feature based methods, Neural Networks and Rule based methods.

Pantic et.al, [6] proposed a multi-detector approach to analyze the spatial changes in the contour of facial components such as the eyes. Geometric-based methods was used to describe the facial actions. They cope up well with variations in skin patterns or dermatoglyphics. But they usually require accurate detection of facial fiducial points, which is difficult when the image has a low-quality or a complex background.

Yang,et.al,[7] proposed that pose, presence or absence of components facial expression, partial or full occlusion, face orientation, lighting conditions are the basic challenges in face detection. Some of the challenges were overcomed but still it maintains a mystery to the researchers. M. Yeasin et al. [8] created discrete hidden Markov models (DHMMs) to recognize the facial expressions.

This paper structured as follows, Section 2 describes the proposed methodology. Section 3 describes the experimental results that were obtained. Section 4 contains final conclusions.

II.PROPOSED METHODOLOGY

The block diagram of the proposed method is shown in figure 1,

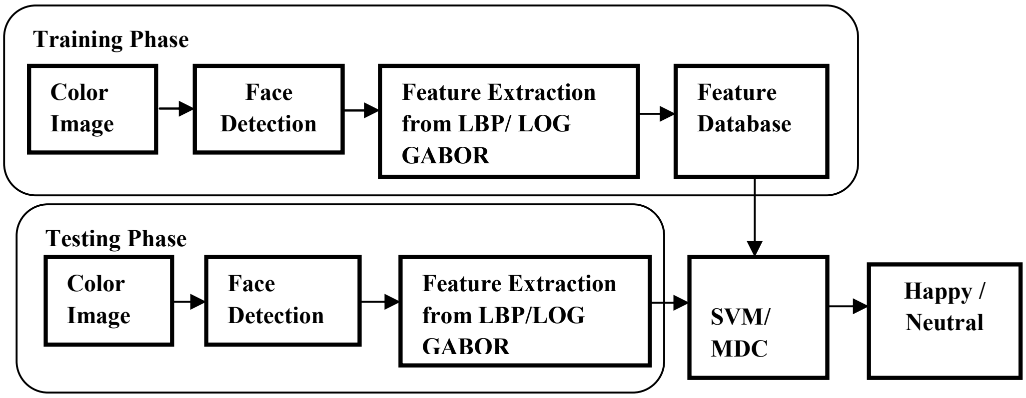


Fig. 1 Block Diagram of Facial Expression Recognition System.

Initially the input color images are grouped into testing images and training images. Training image is normalized into uniform intensity, size and shape. The input image is in RGB color space. It is transformed into YCbCr color space for lighting compensation. The transformed color image is then normalized into same scale size. During this process some amount of noise may occur. This is eliminated using median filter. Median filter is an effective method that suppresses isolated noise without blurring the sharp edges.

Skin tone region is extracted using chrominance value. Using bounding box, the face is alone detected. Then the detected face is extracted by using Local binary pattern operator and Log Gabor transform. For accurate classification, a subset of features are used in the discriminant system instead of using all available variables (features) in the data. The two different facial expressions are classified using support vector machine classifier and the obtained results are compared with the results obtained using Minimum Distance Classifier.

A)Face Detection:

Face detection is the first stage of a facial expression recognition system. Usually, face detection is a two-step procedure: first the whole image is examined to find regions that are identified as “face”. Then localization procedure is done to provide a more accurate estimation of the exact position and scale of the face. Face detection is mostly concerned with roughly finding all the faces in large, complex images. Different structural features like texture, shape and skin color are used. These features are extracted using multi-resolution or derivative filters. Skin color is a powerful cue for detection, as color scene segmentation is computationally fast, while robust to changes in viewpoint, scale, shading. The color-based approach labels each pixel according to its similarity to skin color, and subsequently labels the sub-region as a face if it contains a large blob of skin color pixels.

It is sensitive to illumination ,existence of skin color region and noise. In order to overcome this color transformation is performed. Block diagram of face detection is shown in figure 1.1

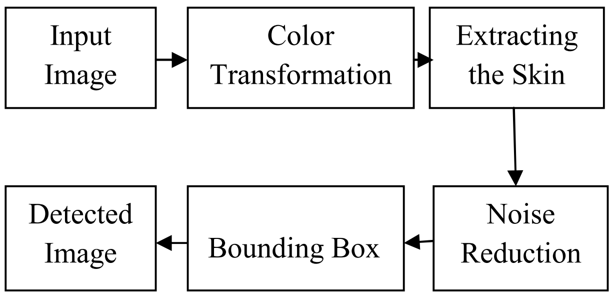


Fig. 1.1 Block Diagram of Face Detection

The aim of the face detection block is to obtain the detected face image with normalized intensity, size and shape. The input color image is typically in RGB format.

i)Color Transformation: RGB components are subjected to the lighting conditions. Thus the face detection may fail if the lighting condition changes. So the image is converted into YCbCr color space for the lighting compensation. In the YCbCr color space, Y contains the luminance information and the chrominance information is in Cb and Cr. Here, the luminance information can be easily de-embedded. The obtained image is then normalized using the formulae

$$new\_Y = \frac{255 * (Y - \min Y)}{(\max Y - \min Y)}$$

ii)Skin Extraction: In the skin color detection process, each pixel was classified as skin or non-skin based on its color components. The detection window for skin color was determined based on the mean and standard deviation of Cb and Cr component.

iii)Noise Reduction: Noise present in the output is removed using the median filter. The median filter is an effective method that suppress the isolated noise without blurring the sharp edges. In median filter, a window slides along the image, and the median intensity value of the pixels within the window becomes the output intensity of the pixel that is being processed.

iv)Bounding Box: Bounding box is applied on each segment with the quantified window size which was selected to meet the size of a face. To merge the separate areas into one area, box-merge algorithm is used which simply merges two or more adjacent square boxes into one. Thus the detected face image after removing the noise is obtained as the output in this face detection phase.

B)Feature Extraction:

Feature extraction phase is a key component of any pattern recognition system. Transforming the input data into the set of features is called feature extraction. Using Local Binary Pattern operators feature extraction is performed. Feature extraction is used to convert pixel data into high level representation. It is used for expression categorization. Feature extraction reduces the dimensionality of input space. It is a special form of dimensionality reduction.

i)Local Binary Pattern Operator: The original LBP operator was introduced by Ojala [9]. It is a powerful method of texture description. LBP is a binary pattern that represents the magnitude relation between the center pixel of a local region and its neighboring
pixels.

The original 3x3 neighborhood is threshold by the value of the center pixel. The values of the pixels in the threshold neighborhood are multiplied by the binomial weights given to the corresponding pixels. Finally, the values of the eight pixels are summed to obtain the LBP number for this neighborhood.

LBP features contains information about the distribution of the local micro-patterns, such as edges, spots, and flat areas over the whole image. All the LBPs of micro patterns in the image are packed into a single histogram. Doing so discards important information concerning spatial relations among the LBPs, even though they may contain information about the image's global structure. To consider such spatial relations, we measure their co-occurrence among multiple LBPs. To improve the performance, an extension of original LBP by considering the co-occurrence of adjacent LBPs, measuring co-occurrence with an auto-correlation matrix generated from multiple LBPs. The proposed feature is robust against variations in illumination, because it only depends on the magnitude relation between a center pixel and its surrounding pixels. The proposed feature is robust against variations in illumination, a feature inherited from the original LBP, and simultaneously retains more detail of image.

ii)Log Gabor Transform: Gabor filters are commonly recognized as one of the best choices for obtaining localized frequency information[1]. They offer the best simultaneous localization of spatial and frequency information. But Gabor functions of arbitrarily wide bandwidth cannot be constructed and it still maintain a reasonably small DC component in the even symmetric filter.

So as an alternative Log Gabor filters are chosen. There are two important Characteristics of Log Gabor functions. Log-Gabor functions, always have no DC component, which enables it to improve the contrast ridges and edges of images. Secondly, the transfer function of the Log Gabor function has an extended tail at the high frequency end, that enables to obtain wide spectral information with localized spatial extent.

This helps to preserve true ridge structures of images. Log- Gabor functions have Gaussian transfer functions when viewed on the logarithmic frequency scale. Using Log Gabor filter function 16 sub bands are obtained for each image. In this paper we have taken 4 scale and 4 orientation. statistical features such as mean and standard deviation were taken to improve the recognition accuracy. Figure 2 shows 16 sub bands obtained from 4 scale and 4 orientation.

C)Support Vector Machine:

Support Vector Machines (SVM) are based on the concept of decision planes that define decision boundaries [10]. SVM perform pattern recognition between two point classes with the help of a surface obtained by using certain points of training data and these points are called support vectors. It is a non-probabilistic binary linear classifier which constructs a hyper plane or a set of hyper plane for the classification. The basic idea behind the SVM classification in the linearly separable data is to choose a hyper plane which gives us the maximum separation of two groups of data. The output of the SVM system is a label that classifies the grid under examination to one of the two basic facial expressions. Coordinate difference between the last and the first frame, is used as an input to the SVM system. In the case of direct facial expression recognition, this system is composed of one two-class SVMs, one for each one of the two basic facial expressions (neutral, happy) to be recognized.

D)Minimum Distance Classifier:

The main idea behind the minimum distance classification method is to calculate the distance of the test point to the classes and then decide the class of the observation with the minimum distance. Pixels that are **"close"** in feature space will be grouped in the same class. The relative distances may change when data are calibrated or atmospherically corrected or rescaled ,in ways that treat different spectral bands differently. If two features have two different units, they must be scaled to provide comparable variance. Otherwise the "distance" will be biased toward the feature with the smallest absolute range.

III.EXPERIMENTAL RESULTS

Facial expression recognition tests were performed using static images from the Position in Global oversight dataset. A total of 120 face images from 24 subjects depicting two different facial expressions: happy and neutral were selected. In the training phase a total of 48 images, 24 from each category were used. In the testing phase 72 images, 24 happy images and 48 neutral images were classified. The images used in the testing set were not included in the training set thus ensuring a person-independent classification of facial expressions. The input color image is transformed into YCbCr color space for illumination compensation. Then the skin region is selected by setting a threshold value. The input image and their corresponding luminance and chrominance output is shown in figure 2

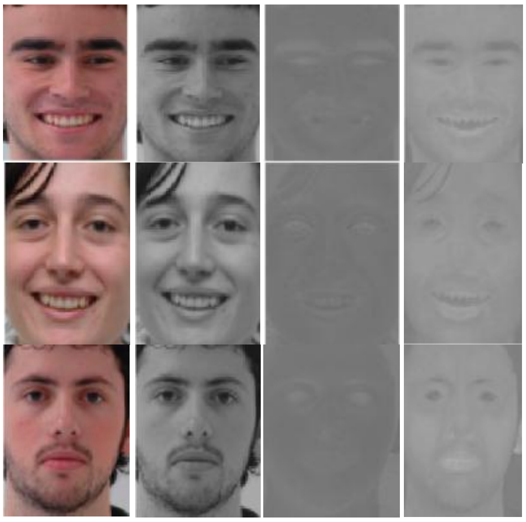
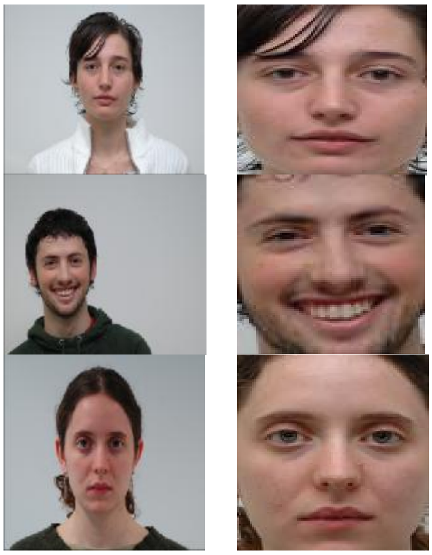


Fig. 2 YCbCr output

Input Image Y Cb Cr

Noise present is removed using median filter. By using a bounding box ,detected face image is obtained. figure 3 depicts the sample input image and the detected face output.



Input Image Detected face

, Fig.3. Detected face images

Orientation / scales

1

2

3

4

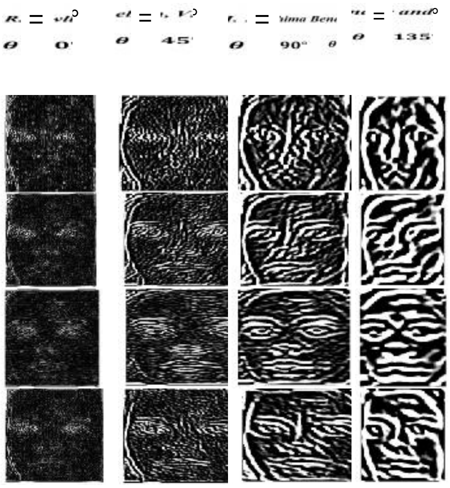


Fig. 4 sub bands obtained using 4 scale and 4 orientation

The features were generated from the detected face output based using Log Gabor filter and LBP operator. These features are robust against variations in illumination. In this paper we have used some features like Mean and Standard deviation. These selected features are then classified using both minimum distance and support vector machine classifier (SVM) for obtaining better results. Figure 5 and figure 6 shows the classification rate for different expressions obtained using support vector machine and minimum distance classifier.

Expression	Number of Images		Recognition Rate (%)	
	Training	Testing	LBP	LOG GABOR
Happy	24	48	100	100
Neutral	24	72	93.75	97.92
Average			**96.87**	**98.96**

Fig. 5 Percentage of classification

On average, the log-Gabor filters classified correctly **98.96**% of cases, whereas the LBP operator method gave the overall rate of **96.87**% using SVM classifier. Log-Gabor filters classified correctly **94.79**% of cases, whereas the LBP operator method gave the overall rate of **92.705**% using MDC classifier. The percentage of correct classifications varied across different facial expressions.

Expression	Number of Images		Recognition Rate(%)	
	Training	Testing	LBP	LOG GABOR
Happy	24	48	95.83	100
Neutral	24	72	85.98	89.58
Average			**92.705**	**94.79**

Fig. 6 Percentage of classification

IV.CONCLUSIONS

In this paper, two different methods of feature extraction for facial expression recognition system: the log-Gabor filters and LBP operator was studied. The proposed feature extraction methods are robust against variations in illumination. To improve their performance, co-occurrence features were considered in addition. The results showed that the log-Gabor method outperformed the LBP method producing the largest improvement in the classification accuracy and in the discrimination between different facial expressions. The LBP operator needs less memory and has less features than log-Gabor filters. Though the log-Gabor filters method have better recognition ratios than LBP operator, its not good to some expressions and also the computational load is complex and time consuming. These selected features are then classified using support vector machine classifier (SVM) and Minimum Distance classifier for obtaining better results. On average, the log-Gabor filters classified correctly **98.96**% of cases, whereas the LBP operator method gave the overall rate of **96.87**% using SVM classifier. Log-Gabor filters classified correctly **94.79**% of cases, whereas the LBP operator method gave the overall rate of **92.705**% using MDC classifier. The percentage of correct classifications varied across different facial expressions.

References

[1] Seyed Mehdi Lajevardi and Zahir M. Hussain "Facial Expression Recognition Using Log-Gabor Filters and Local Binary Pattern Operators" International Conference on communication, computer and power (ICCCP'09) Muscat feb.15-18, 2009.

[2] Chen,et.al,"Hybrid-boost learning for multi-pose face detection and facial expression recognition and Pattern Recognition", vol. 41, no. 3, pp. 1173,2008.

[3] M.H.Yang et.al , "Detecting faces in images: A Survey", IEEE Trans.on pattern analysis and machine intell.,24(1),2002.

[4] I. Cohen et.al, "Facial expression recognition from video sequences: temporal and static modeling, Computer Vision and Image Understanding" 91, 160–187,2003.

[5] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: the state of the art", IEEE Trans.

Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1424–1445, Dec. 2000.

[6] M. Pantic and L. J. M. Rothkrantz. “Facial action recognition for facial expression analysis from static face images”. IEEE Trans. SMC, Part B: Cybernetics, 34(3):1449–1461, 2004.

[7] Yang, M. H., et.al, “Detecting faces in images: A survey,IEEE Trans. on Pattern Analysis and Machine Intelligence”. vol. 24, no. 1, pp. 34–58,2002.

[8] M. Yeasin, et.al. “Recognition of Facial Expressions and Measurement of Levels of Interest From Video”. IEEE Trans. Multimedia,8,3,2006.

[9] Ojala T et.al, “A comparative study of texture measures with classification based on feature distributions”, Pattern Recognition, Vol 29, No. 1, pp.5158-5159, 1996.

[10] C.W. Hsu and C. J. Lin, “A comparison of Methods for multiclass Support Vector Machines”, IEEE Transactions on Neural Networks, vol. 13, no. 2, pp. 415-425, 2002.

[11] Bourel et.al "Robust Facial Expression Recognition Using a State-Based Model of Spatially-Localised Facial Dynamics", Proc. Fifth IEEE International Conf. Automatic Face and Gesture Recognition, pp. 106-111, 2002.

[12] D. Field. “Relations between the statistics of natural images and the response profiles of cortical cells”. J. Opt. Soc. Am.,A(4):2379 – 2394, 1987.

[13] Hanchuan Peng et.al “Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp.1226-1238, 2005.