Face Emotions Recognition using Compounded Local Binary Pattern (CLBP) Feature Descriptor with Support Vector Machine (SVM)

Farah Jamal Ansari¹, Vaseem Durrani²

¹University Polytechnic, Faculty of Engineering & Technology, Jamia Millia Islamia, New Delhi- 110025, India
²Aedifico Pvt. Ltd, New Delhi-110092

Abstract— An automatic face feature detection problem from a static front pose image is discussed in this paper. It deals with the classification and recognition of face emotions or expressions and also about the different signs of moods of different people. Support Vector machine (SVM) has been used to classify and identify the expressions of a given face into five categories. The five categories are neutral, sad, happy, disgust and angry. In this paper, we propose a reliable and robust facial feature vector exploring the (CLBP) Compound Local Binary Pattern for recognition of face emotion which overcomes many of the shortcomings of LBP(Local Binary Pattern) To implement the new improved feature descriptor the original LBP is joined with the add-on P bits in our proposed method that extracts both the magnitude and the sign of the differences between the centre and neighboring pixel values. JAFFE facial expression database is taken into consideration to perform the experiments. 100% accuracy has been obtained for the training set and 96.26% accuracy is obtained for the testing set and a cohn kanada database is taken into consideration in which we got 97.61% recognition rate for the facial expression.

Keywords- Facial Expression Recognition, Facial features, feature extraction, Space Vector Machine, CLBP.

I. INTRODUCTION

With the advancements in technology; there is an emergence of the era where humans and intelligent robots will share the same universe. Work and live hand in hand. Human Computer Interaction (HCI) area plays a vital aspect to resolve unavailability of unbiased understanding in interaction among humans and machines. Human Computer interaction can become furthermore efficient and worthy if machine predicts about human emotional condition at any instance and also the mood or emotion of a person out of any given picture based on various face expressions of the person. Mehrabian [1] has found that human communication information is communicated by the verbal for about 7%, and vocal is about 38% and facial expressions is about 55%.

Therefore, facial expressions are one of the most unique features and useful data for understanding of emotional front of the human conversation face to face. It is important to extract main core facial features for categorizing the various facial expressions into different categories. This contributes in detecting proper and distinguished expressions. In vision community, recognizing and classifying human facial expression is a big challenge for the development of automatic expressions recognition system. Recently, lots of researches are concentrated on recognition of human facial expressions by computer [2-6].

In previous years, using computer software for Expression of faces and recognition of emotions and utilizing the data related to it in human computer interaction has developed a major research interest because of which it has given a increase in number of automatic methods or approaches for recognizing or detecting facial expressions in pictures and videos [7-12]. This paper explained a robust method to the problem of extracting descriptors of faces from a fixed frontally posed picture and its identification and recognition of emotion of faces and person’s at present mood. Support vector Machine is used as a classifier to identify the expressions of an input face into five categories. The five categories include sad, neutral, happy, angry and disgust. Face segmentation has been done by using image processing morphological operations like erosion, dilation, and reconstruction. In the proposed technique we have used Compounded LBP for feature extraction with SVM as a classifier. Experiments have been performed on the JAFFE and Cohn kanada facial expression database.

II. LITERATURE SURVEY OF EXISTING METHODS

Currently, the study on development of automatic recognition system for face expressions has been a matter of research. The meticulousness of facial expression recognition has the core emphasis on exact facial feature component extraction.
It primarily comprises 3 forms of data namely shape, texture and permutation of both shape and texture info. To find the permutation of global data, local feature data and shape data for creating a feature vector, Fenget al. [15] has used LBP and AAM. They have used adjacent zone with biased chi-sq information for classification of emotions. With the help of AAM the feature point localization is completed and Centre of mouth and eyes is gained based on them. To excerpt facial features for instance mouth corners and center, eye corners and center, nose corner and chin and cheek border etc., Mauricio Hess and G. Martinez [16] have used SUSAN algorithm. Gengtaozhou et al. [17] utilized the selective feature extraction technique where emotions are usually categorized into 3 kinds as per the to the deforming of mouth as 1) fear, happy 2) Disgust, anger, sad 3) surprise. Certain if then rules are applied to further categorize separate group expressions. For facial feature localization Gabor wavelet filter has used Jun Ou et al. [18], 28 facial feature facts and PCA for feature extraction and KNN for classification of expression. Md. Zia Uddin ,T.S. Kim and J.J. Lee [19] used EICA (Enhanced Independent Component Analysis) to tweak out nearby independent component features which are more over categorized by (Fisher linear Discriminant Analysis) FLDA. Subsequently discrete HMM is used to carve unrelated Expressions of faces. The outcomes of feature descriptor extraction of numerous well known conventional technique (PCA-FLDA, PCA, EICA and ICA) in combination with similar Hidden Markov Model structure has been likened and comparative study in terms of the (RR) recognition rate. PCA is an unsubstantiated learning technique which is made to take beneficial characteristics and second order arithmetical technique is used for originating orthogonal bases holding the supreme in consistency and is also used for dimension reduction.V. Gamathi et al.[20] used constant LBP histogram method for feature descriptor extraction and M-ANFIS (Multiple Adaptive Neuro Fuzzy Inference system) for emotions recognition. For identifying basic expression, Hadi Seyedarabi et al.[21] used facial expression recognition system. Cross correlation based optical flow technique is used for extracting facial feature vectors. . GRS Murthy and R.S. Jadon[22] have used the improved PCA for Eigen face modernization technique in emotion recognition. They have parted the training set of JAFFE and Cohn kanade databases into six diverse dividers. The eigen space is made for every one class and after that the picture is rebuilt. A resemblance measure used for associating unique and rebuilt picture is Mean square error. Zhengyou Zhang et al. [23] represented a FER system.

In this the use of two varieties of features extracted from face images for recognizing facial expression has been likened. The two methods utilized for feature extraction are a set of fiducial point geometric positions and multi-scale and multi orientation gabor coefficients extracted from the face image at the fiducial. These are specified to support vector machine classifier discretely or together and up shots were associated.

III. DATABASE USED

For research required data is collected from JAFFE (Japanese Female Facial Expression) database for support vector machine teaching and testing. This database comprises of 213 pictures of seven face expressions (six basic facial expressions + one neutral) which are of 10 diverse Japanese females. 60 Japanese subjects have graded every single image on six sentiment adjectives.

Fig. 1 facial expressions sample of person YM

Fig. 2 facial expressions sample of cohn kannada database

IV. (CLBP) COMPOUNDED LOCAL BINARY PATTERN.
A. Encoding Method- CLBP

Unique Local Binary Pattern encryption subject deliberates exclusively the symbol of the distinction among two grey values and so, it generally miss the mark to acquire binary codes in line with the feel stuff of a neighborhood area. Being compelled by this, we have a propensity to plan CLBP, subordinate elongation of the 1st operator of LBP that allots a 2P-bit code to the mid image component supported the neighborhood grey values involving neighbors (P).
Disparate from the LBP operator that pays 1 bit for every neighbor to signify exclusively the symbol of the distinction among the mid and as well the parallel neighbor grey values, the planned approaches use two bits for every single neighbor in order to code the symbol thus far as the magnitude data of the distinction among the mid and as well the neighbor grey values. At this point, the main bit signifies the symbol of the distinction among the mid and as well the parallel neighbor grey values just similar to the basic LBP encryption. The opposed bit is working to code the magnitude of the distinction with relevance a threshold price, that is the average magnitude $M_{avg}$ of the distinction among the mid and as well the neighbor grey values within the native neighborhood of concentration. If the magnitude of the distinction among the mid and as well the parallel neighbor is superior to the edge $M_{avg}$ then the CLBP operator sets this bit to 1. Else, it’s set to zero. Therefore, meter s(x) of equation a pair of is swapped by the following function:

$$s(i_p, i_c) = \begin{cases} 
00 & |i_p - i_c| < |M_{avg}| \\
01 & |i_p - i_c| > |M_{avg}| \\
10 & i_p - i_c \geq 0, \quad |i_p - i_c| \geq |M_{avg}| \\
11 & \text{otherwise}
\end{cases}$$

(1)

At this point, ic is that the grey worth of the middle image element, informatics is that the grey price of a p neighbor, here $M_{avg}$ is that the average magnitude of the distinction among informatics and ic inside the native neighborhood. Simple Compound-LBP operator as shown in Fig. 2.

![Figure 3.Basic CLBP operator, wherein, the Compound-LBP code = 1010101011011111 for pixel C.](image)

After looking out in Fig3, it's clear that the planned CLBP approach separates the neighbors within the south-east, east, and north-east orders as they require greater grey values as compared to the opposed neighbors and therefore, creates a designing line with native texture stuff.

### B. Extraction of Sub-CLBP descriptors

A 3x3 neighborhood is taken and a proposed Compound LBP technique computes a picture by operative on the eight neighbors round the central element and distribution a 16-bits code to its element. As a 16-bit codes are accustomed label the image pixels, the amount of attainable binary patterns is 216. In order to diminish the amount of options, He and Cercione [27] work to consider lesser neighbors numbers while generating the binary (0 1) patterns. Therefore, the length of the descriptor of feature is often diminished by discarding some extent of neighborhood data. Here, we’ve got conferred a distinct method wherever all the Compound LBP binary patterns are more divided into two sub-Compound LBP patterns. All sub-Compound LBP pattern is obtained by joining the bit values adore P/2 neighbors, wherever P is that the variety of neighbors. Formally, in an exceedingly native neighborhood, the 2 sub-Compound LBP patterns are fashioned by joining the corresponding values of the bit sequence (3, 4, 7, 8, ..., 2P-1, 2P) and (1, 2, 5, 6, ..., 2P-3, 2P-2), severally of the 2P-bit original Compound LBP code.

In alternative way, a 16-bits Compound LBP pattern is divided into parts as two 8-bits sub-Compound LBP patterns, wherever primary sub-CLBP1 is generated by combining the values adore the neighbors within the north, east, south, and west directions, severally and also second sub-Compound LBP pattern sub-CLBP2 is extracted by combining the bit values adore the neighbors within the south-east, north-east, south-west, and north-west directions, severally. This technique diminishes the amount of attainable patterns considerably, which ends in an exceedingly total of twenty eight different and separate sub-Compound LBP patterns. The method is explained in Fig4. The 2 sub-CLBP patterns are treated as distinguished binary codes and concatenated throughout the descriptors of feature generation.

![Fig.4. Extraction of two sub-CLBP binary pattern 10111010 and 1111010 from binary original CLBP code 1011111110101010.](image)
C. Compounded LBP Feature Descriptor

The Compounded LBP operator is applied on all the pixels of image and the 16-bit Compounded LBP patterns is separated to form sub-Compounded LBP patterns. Like this we form two 8-bit binary code for every associated element of the image. This results into two encoded images for two sub-compounded LBP patterns. For these two encoded images histograms are generated and concatenated to make a single histogram, the Compounded LBP histogram. This is used as the feature descriptor for the facial expression image. Figure 5 demonstrates the Compounded LBP task for generating histogram using a sample emotion image.

Then finally, all the regions bar graphs are concatenated to get the extended Compounded LBP histogram. For the robust recognition method, this bar graph inclusion and embedding is employed for the facial feature vector. The extended process of histogram extraction is shown in Fig. 6.

D. (SVM) Support Vector Machine as a Classifier

SVM is basically for binary classification. And after certain modification is can work for multi class. Firstly the binary SVM is discussed. Thereafter, we tend to discuss however this method is often more elongated or modified to deal issues with general multi-class classification.

Binary Classification

SVM is the part of most margin classifiers category. They mostly perform pattern or object recognition between two categories. They notice a choice surface that has most distance with the nearest points within the coaching set. Those termed as support vectors. we are tend to taking a coaching set of points x(i) \{IR\}n, wherever every purpose x(i) belongs to at least one of available categories known by the label y(i) .we are forward linearly dissociable data, the goal of most of the margin classification is to differentiate the available categories by a hyper-plane specified the space to the support vectors is maximized. Hyper plane basically are the best known and understood as optimum separating hyper plane (OSH).The form of OSH [24]:

\[ f(x)= \sum_{i=1}^{L} \alpha_i y_i x_i^T x + b \]  

(2)
Coefficients $b$ and $\alpha_i$ in Eq. (1) are the quadratic problem solutions. Specific latest data point classification is done by calculating the sign of the right side of Eq. (3). We will use [24]

$$d(x) = \frac{\sum_{i=1}^{N} a_i y_i \phi(x_i) X^T x + b}{\| \sum_{i=1}^{N} a_i y_i \phi(x_i) \|}$$

(3)

in order to complete multi-class identification or classification. The sign of $d$ is the classification result for $x$. $|d|$ the magnitude of $d$ is the distance from the hyperplane. The larger $|d|$ the magnitude of the $d$ means the more far away the point is from the surface, the more accurate result of the classification is.

The complete implementations are often apprehended to the nonlinear separating surfaces [24]. Each purpose within the input area has been mapped to some extent $z=\Phi(x)$ of a advanced dimensional area that is understood because the feature area. In the feature space the information are separated by a hyper-plane. The very important aspect during this structure is that the mapping $\Phi(.)$ is subject to the condition that the real of two points within the characteristic space $\Phi(x) \Phi(y)$ are often termed as a $K(x,y)$ kernel function.

Equation for decision surface is given as [24]:

$$f(x) = \sum_{i=1}^{N} y_i a_i K(X_i, x) + b_i$$

(4)

Multi-class classification

There are two basic ways to solve the multi category issues with SVMs. These are:

1. SVMs are trained within the one-vs.-all classification approach. Here every single category separates one category from the remained categories by the SVM [5, 17].
2. The machines are trained within the pair-wise approach. Every single SVM segregates a try of categories. The group wise classifiers are organized in trees wherever all SVM is described by tree node. A bottom-up tree was projected in [16] for recognition of objects. It was applied to recognition of faces in [7]. Recently, a top-down tree structure is printed in [15].

No notional explanation for the two above mentioned ways with reference to performance of the classification. With regards to coaching effort, the one-vs-all method is a lot of most well-liked as a result of solely SVMs is trained as compared with SVMs within the second method. The run-time complexness of each the ways is sort of alike.

Recently, experiments are performed on person recognition to point out matching performance of classification for the two ways [13]. Since the categories number in recognition of faces are generally huge enough, we’ve got opted the one-vs-all technique rather than pair-wise, as a result of the amount of SVMs used is linear with the amount of categories.

V. EXPERIMENT AND RESULTS

We have considered a JAFFE database for the facial expression recognition. And created a complete GUI in the MATLAB software tool to get the complete analysis as shown in the figure 7. The overall recognition rate that we have achieved is 100 percent in approximately 100 samples as the testing.

Fig.7: GUI screenshot

Fig. 8: Confusion matrix

For one instance figure 8 displays the confusion matrix of overall recognition of all five facial expression that we have considered like anger, disgust, happy, neutral and sad. Here we have considered the small database and achieved the 100% recognition rate. But as the database increases the recognition rate tends to reduce.
International Journal of Emerging Technology and Advanced Engineering

Table 1:
Recognition Rate of two databases used

<table>
<thead>
<tr>
<th>Database Taken</th>
<th>Recognition Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAFFE</td>
<td>96.26</td>
</tr>
<tr>
<td>Cohn Kanada</td>
<td>97.61</td>
</tr>
</tbody>
</table>

VI. CONCLUSION

In this paper a compounded LBP technique which is the modified LBP technique has been proposed to achieve the better performance. The experiment results obtained have shown the robustness of the proposed technique on two databases JAFFE with 95.26% recognition rate and Cohn Kanada with 97.61% recognition rate.

REFERENCES

[23] Nectarios Rose, —Facial Expression Classification using Gabor and Log-Gabor Filters, Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FG’06) IEEE computer society