International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017

Face Emotion Recognition Using Combined Sift, LBP and Local Phase Quantization Features

X.Arputha Rathina,
Associate Professor, B.S.A Crescent University Chennai

ABSTRACT:

Recognizing facial expression has remained a challenging task in computer vision. Deriving effective facial expression recognition is an important step for successful human-computer interaction systems. This paper describes a novel approach towards facial expression recognition task. Subsequently, we extract the features from Image using both Local Binary Pattern (LBP) and Local Phase Quantization (LPQ). Then the results from both texture descriptors are tested on the Facial Expression Recognition. Occlusion is a big challenge for facial expression recognition (FER) in real-world situations. The proposed here, two different methods of feature extraction for facial expression recognition from occluded images. The local binary pattern operator (LBP) and local Phase Quantization (LPQ) are used for feature extraction. Six basic facial expressions are considered. We consider four types of very frequently occurred occlusions in real-world situations, the eyes/mouth and upper/lower face region occlusion. The experiments show that the proposed approach is robust to partial occlusion of the face.

1. INTRODUCTION

Facial expression recognition (FER) has become a hot research topic of human-computer interaction (HCI) and drawn a lot of attention due to its great potential in multimedia applications, e.g. digital entertainment, customer service, driver monitoring and so on. HCI would become more friendly and natural if computers are able to recognize affects as human beings, which can benefit from solving FER problems.FER aims to analyze and classify a given facial image into one of the six commonly used emotion types where the six emotion categories are angry, disgust, fear, happy, sad and surprise. Numerous algorithms of FER have been proposed in the literatures during the past several years, including expression recognition from frontal and non-frontal facial images. Comparing to frontal FER, non-frontal FER is more challenging and more applicable in real scenarios.

However, only a small part of algorithms among the proposed various methods address this challenging issue. For both frontal and non-frontal FER problems, a general recognition framework appeared in most of previous works can be divided into two major steps, one is the feature extraction and the other is classifier construction.

To extract the facial features, various image features are employed in the previous papers, such as local binary pattern (LBP) local phase quantization , histograms of oriented gradients and scale-invariant feature transform (SIFT). Among the various facial features, SIFT has demonstrated promising performance due to its robust property to image scaling, rotation, occlusion and illumination difference.

2. METHODOLOGIES

- Input face image
- SIFT based invariant Features
- Local Binary Pattern(LBP)
- Local Phase Quantization(LPQ)
- Key point Matching and indexing.

2.1 Input face image

It is process to extract face regions from input image which has normalized intensity and uniform in size. The appearance features are extracted from detected face part which describes changes of face such as furrows and wrinkles (skin texture). Here, MATLAB vision toolbox will be utilized to detect the face from input .It returns the face boundary vector [X_{min}, Y_{min}, Width and Height] to crop the desired region for further texture analysis.

2.2 Scale invariant detector

The first step is to turn the input image into an integral image. This is done by making each pixel equal to the entire sum of all pixels above and to the left of the concerned pixel which is represented in the below figure.

1	1	1
1	1	1
1	1	1

Input image

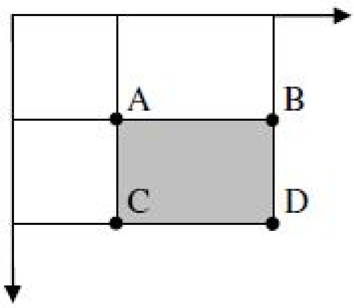
1	2	3
2	4	6
3	6	9

Integral image

Fig 1 Scale Invariant Detector

This allows for the calculation of the sum of all pixels inside any given rectangle using only four values.

These values are the pixels in the integral image that coincide with the corners of the rectangle in the input image. This is demonstrated in figure 3.2



Sum of grey rectangle = D - (B + C) + A

Fig 2 Gray Rectangle

Since both rectangle B and C include rectangle A the sum of A has to be added to the calculation. It has now been demonstrated how the sum of pixels within rectangles of arbitrary size can be calculated in constant time. The Viola-Jones face detector analyzes a given sub-window using features consisting of two or more rectangles.

Each feature results in a single value which is calculated by subtracting the sum of the white rectangle(s) from the sum of the black rectangle(s).It is empirically found that a detector with a base resolution of 24*24 pixels gives satisfactory results. When allowing for all possible sizes and positions of the features in Figure 4 a total of approximately 160.000 different features can then be constructed. Thus, the amount of possible features vastly outnumbers the 576 pixels contained in the detector at base resolution.

These features may seem overly simple to perform such an advanced task as face detection, but what the features lack in complexity they most certainly have in computational efficiency. The hope being that some features will yield large values when on top of a face.

Of course operations could also be carried out directly on the raw pixels, but the variation due to different pose and individual characteristics would be expected to hamper this approach.

2.3 Local Binary Pattern

Local Binary Patterns (LBP) is one of the most used methods in face recognition. To improve the recognition rate and robustness, several methods using LBP, have been proposed. Improved Local Binary Pattern (ILBP) is an improvement of LBP which compare all the pixels (including the center pixel) with the mean of all the pixels in the kernel to improve the robustness against the illumination variation. For the purpose of retaining the spatial and gradient information, an extended version of Local Binary Patterns (ELBP) that encodes the gradient magnitude image in addition to the original image was propose to represent the velocity of local variation. Local Gabor Binary Pattern (LGBP) is another representation approach based on multi-resolution spatial histogram combining local intensity distribution with the spatial information via introducing the Gabor wavelets into the LBP as the image pre-processing. The LBP code is calculated in the belowfigure..

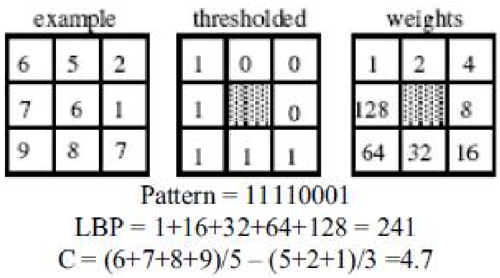


Fig 3 Calculating original LBP code

Local feature based approaches have got great success in object detection and recognition in recent years. LBP encode local primitives including different types of curved edges, spots, flat areas, etc. The advantage of LBP was invariant to monotonic changes in gray scale. So LBP is widely used in face recognition, pedestrian detection and many other computer vision applications.

The derivation of the LBP follows that represented by Ojala et al. (2002). Let us define textureas the joint distribution of the gray levels of *P* +1 (*P* > 0) image pixels:

$$T = t(g_c, g_0, \cdots, g_{p-1})$$

where *c g* corresponds to the gray value of the center pixel of a local neighbourhood $g_p\,(p = 0, \cdots, P-1)$ Correspond to the gray values of *P* equally spaced pixels on a circle of radius *R*(*R* > 0) that form a circularly symmetric set of neighbours illustrates three circularly symmetric neighbour sets for different values of *P* and *R* .

Without losing information, *c g* can be subtracted from *p*

$$T = t(g_c)t(g_0 - g_c, \cdots, g_{p-1} - g_c)$$

Since ( ) *c t g* describes the overall luminance of an image, which is unrelated to local image texture, it can be ignored:

$$T \approx t(g_0 - g_c, \cdots, g_{p-1} - g_c)$$

A binomial weight 2P is assigned to each sign ( ) p c s g − g , transforming the differences in a neighborhood into a unique LBP code:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^P$$

The square root of the pixels is taken. Then, the first order gradients are computed. The gradient magnitude at each pixel is then computed as,

$$\omega_{x,y} = \sqrt{I_x^2 + I_y^2}$$

where *Ix* and *Iy* are the first-order derivatives in the *x* and*y* directions. *ωx,y* is then used to weigh the LBP code. The stronger the pixel contrast, the larger the weight assigned to the pixel LBP code. In this way, if a LBP code covers both sides of a strong edge, its gradient magnitude will be much larger and by voting this into the bin of the LBP code, we take into account if the pattern in the local area is of a strong contrast.

Thus, the resulting feature will contain both edge and texture information in a single representation. The value of the *i th* weighted LBP bin of a *M* × *N* block is as follows:

$$h_{lbp}(i) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta(LBP_{x,y}, i)$$

$$\delta(m,n) = \begin{cases} 1, & m = n \\ 0, & \text{otherwise} \end{cases}$$

The RLBP histogram is created from (6) as follows

$$h_{rlbp}(i) = h_{lbp}(i) + h_{lbp}(2^B - 1 - i), \quad 0 \le i < 2^{B-1}$$

where *hrlbp(i)* is the *i th* RLBP bin value. To mitigate the RLBP issue in Fig. 2, consider the absolute difference between the bins of a LBP code and its complement to form difference of LBP (DLBP) histogram as follows

$$h_{dlbp}(i) = |h_{lbp}(i) - h_{lbp}(2^B - 1 - i)|, \quad 0 \le i < 2^{B-1}$$

where *hdlbp(i)* is the *i th* DLBP bin value. The number of DLBP bins is 128 for B = 8. Using uniform codes, it is reduced to 30. For blocks that contain structures with both LBP codes and their complements, DLBP assigns small values to the mapped bins. It differentiates these structures from those having no complement codes within the block. The 2 histogram features, RLBP and DLBP, are concatenated to form Discriminative Robust LBP (DRLBP) as follows:

$$h_{drlbp}(j) = \begin{cases} h_{rlbp}(j), & 0 \le j < 2^{B-1} \\ h_{dlbp}(j - 2^{B-1}), & 2^{B-1} \le j < 2^B \end{cases}$$

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017

2.4 Local phase quantization

The face image is first labeled with the LPQ operator. Then, the label image is divided into non-overlapping rectangular regions of equal size and a histogram of labels is computed independently within each region. Finally, the histograms from different regions are concatenated to build a global description of the face. Using LPQ operator to Extract Expression Features In the experiment, LPQ features are first calculated for each expression image. Then each LPQ image is divided into 5×3 non overlapping blocks of the same size. LPQ histogram rather than LPQ itself of each block is used as the facial expression features. In each region, a histogram of labels is computed independently. Then we concatenate with all the histograms from different regions and these are a global description of the facial expression images which is represented in the belowfigure.

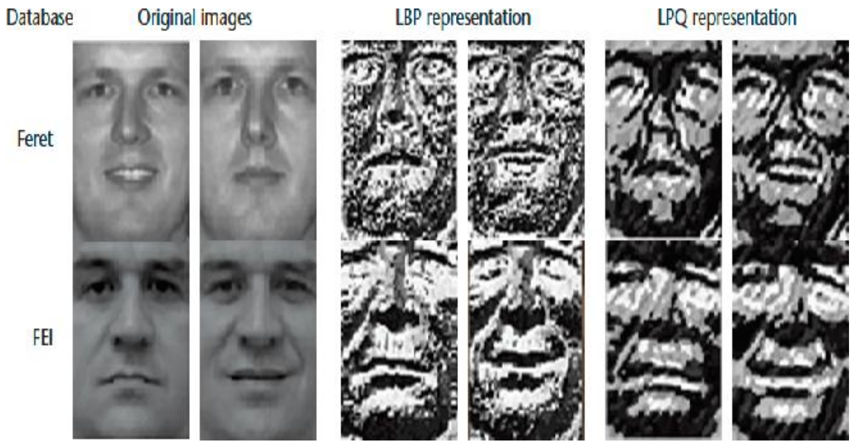


Fig 4 Local Phase Quantization

2.5 Feature matching and indexing

Indexing consists of storing SIFT keys and identifying matching keys from the new image. The best candidate match for each key point is found by identifying its nearest neighbor in the database of key points from training images. The nearest neighbors are defined as the key points with minimum Euclidean distance from the given descriptor vector. The probability that a match is correct can be determined by taking the ratio of distance from the closest neighbor to the distance of the second closest.

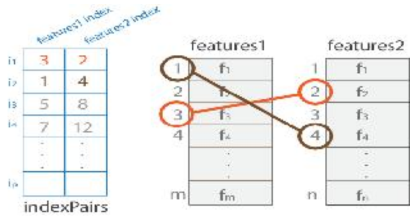


Fig 5 Feature Matching

Indices of corresponding features between the two input feature sets, returned as a *P*-by-2 matrix of *P* number of indices is described in the above. Each index pair corresponds to a matched feature between the features1 and features 2 inputs. The first element indexes the feature in features1. The second element indexes the matching feature in features2.

4. JUSTIFICATION TO THE PROPOSED METHODS

4.1. Face Recognition with PCA

Independent Component Analysis (ICA) to extract the face texture features. Two different frameworks of ICA are adopted to compare with PCA for the recognition performances by using three different classification techniques. Framework I observed images as random variables and the pixels as outcomes while framework II treated pixels as random variables and the images as outcome. This will able to show that ICA framework II yields the best performance for identifying faces and it is able to provide both False Acceptance Rate (FAR) and False Rejection Rate (FRR) as low as 1%.

Therefore face recognition offers promising future for medium-security access control system. An important issue in face recognition is to extract face features that can discriminate an individual from the other. There are two popular approaches to face recognition. One approach transforms face images into specific transformation domains. Among the works that appear in the literature are Eigen face, Gabor filters, Fourier Transform, and wavelets. Another approach is to extract principal lines and creases from the face. However, this method is not easy because it is sometimes difficult to extract the line structures that can discriminate every individual well

Drawbacks

- Poor discriminatory power
- High computational load

4.2 Face Recognition based on DCT Feature Extraction

In the field of image processing and recognition, discrete cosine transforms (DCT) and linear discrimination is two widely used techniques. Based on them, we present a new face and face recognition approach in this paper. It first uses a two dimensional seperability judgment to select the DCT frequency bands with favourable linear seperability. Then from the selected bands, it extracts the linear discriminative features by an improved Fisher face method and performs the classification by the nearest neighbour classifier. We detailed analyze theoretical advantages of our approach in feature extraction.

The experiments on face databases and face database demonstrate that compared to the state-of-the-art linear discrimination methods, our approach obtains better classification performance. It can significantly improve the recognition rates for face and face data and effectively reduce the dimension of feature space. Frequency domain analysis is a commonly used image processing and recognition technique. During the past years, some work has been done to extract the frequency- domain features for image recognition. Li *et al.* extract Fourier range and angle features to identify the face image. Lai *et al.* use holistic Fourier invariant features to recognize the facial image.

Another spectral feature generated from singular value decomposition (SVD) is used by some researchers. However, Tian *et al.* indicate that this feature does not contain adequate information for face recognition. In, Hafed and Levine extract discrete cosine transform (DCT) feature for face recognition. They point out that DCT obtains the near-optimal performance of Karhunen–Loeve (KL) transform in facial information compression.

And the performance of DCT is superior to those of discrete Fourier transform (FT) and other conventional transforms Linear discrimination technique is thus, important in extracting effective discriminative features and reducing dimensionality of image. This technique usually needs much less computational cost than nonlinear recognition techniques like neural network. So far many linear discrimination methods have been proposed for use in image recognition.

Drawbacks

- In appearance based methods, less accurate of features description because of whole image consideration
- In geometric based methods, the geometric features like distance between eyes, face length and width, etc., are considered which not provides optimal results.

4.3 Face Recognition with LBP

Local Binary Patterns (LBP) is one of the most used methods in face recognition. To improve the recognition rate and robustness, several methods using LBP, have been proposed. Improved Local Binary

Pattern (ILBP) is an improvement of LBP which compare all the pixels (including the center pixel) with the mean of all the pixels in the kernel to improve the robustness against the illumination variation. For the purpose of retaining the spatial and gradient information, an extended version of Local Binary Patterns (ELBP) that encodes the gradient magnitude image in addition to the original image was propose to represent the velocity of local variation. Local Gabor Binary Pattern (LGBP) is another representation approach based on multi-resolution spatial histogram combining local intensity distribution with the spatial information via introducing the Gabor wavelets into the LBP as the image pre-processing; therefore, it is robust to noise and local image transformations.

Local feature based approaches have got great success in object detection and recognition in recent years. The original LBP descriptor was proposed by Ojala et al. and was proved a powerful means for texture analysis. LBP encode local primitives including different types of curved edges, spots, flat areas, etc. The advantage of LBP was invariant to monotonic changes in gray scale. So LBP is widely used in face recognition, pedestrian detection and many other computer vision applications. The basic LBP operator assigns a label to every pixel of an image by thresholding the 3 x 3-neighborhood and considering the results as a binary number. Then the histogram of labels can be used as descriptor of local regions.

5. EXPERIMENTAL RESULTS

5.1 ARCHITECTURE DIAGRAM

The face image is given to the face detection part where the facial features like eyes, mouth and nose. The SIFT algorithm is applied to detect the emotions of facial features. LBP and LPQ features are applied to get the desired emotion recognition. The combined features are compared with the features of database and key point matching is found to recognize the emotions. The design of emotion recognition is clearly depicted in the below figure.

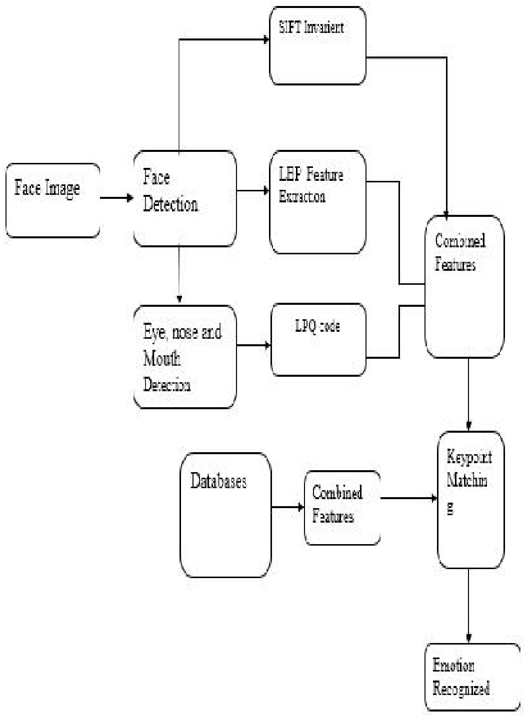


Fig 6: ARCHITECTURE DIAGRAM

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017

Fig 7: Face recognised o/p

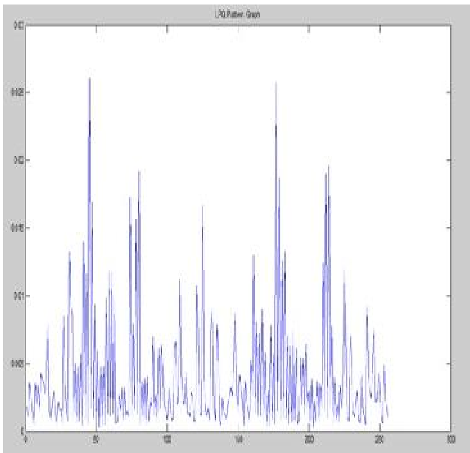
Face detection is usually considered as an important preprocessing step for a recognition operation that allows a reduction of the noise within an image. The appearance of input image is detected in a rectangular thin line. The face regions are extracted from input image which has normalized intensity and uniform in size. Which describes changes of face such as furrows and wrinkles (skin texture)? The basic local binary pattern operator introduced and it is based on the assumption that texture has locally two complementary aspects, a pattern and its strength.

The LBP was proposed as a two-level version of the texture unit to describe the local textural patterns. The original version of the local binary pattern operator works in a 3 × 3 pixel block of an image. The pixels in this block are threshold by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighborhood consists of 8 pixels, a total of 28 = 256 different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood. From this local binary pattern feature, the texture of the input image is detected.

Fig 8: local binary pattern

The image is divided into non-overlapping rectangular regions of equal size and a histogram of labels is computed inde-pendently within each region. Finally, the histograms from different regions are concatenated to build a global description of the face. Using LPQ operator to Extract Expression Features In the experiment, LPQ features are first calculated for each expression image. Then each LPQ image is divided into 5×3 non overlapping blocks of the same size. LPQ histogram rather than LPQ itself of each block is used as the facial expression features. The LPQ pattern is shown below.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 4, Issue 8
August 2017



In each region, a histogram of labels is computed independently. Then we concatenate with all the histograms from different regions and these are a global description of the facial expression images.

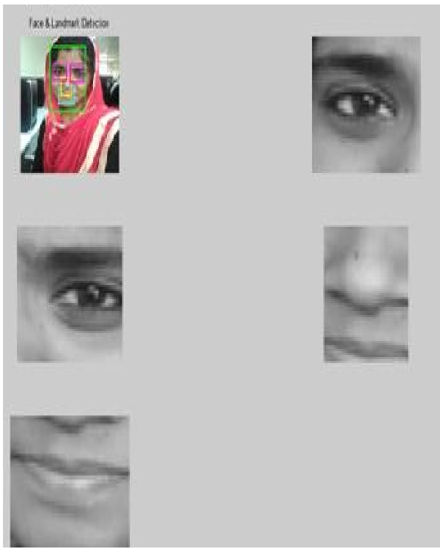


Fig 9: Face and Landmarks

The best candidate match for each key point is found by identifying its nearest neighbor in the database of key points from training images. The emotion of the particular facial expression is recognized with the help of combined features of LBP and LPQ.

6. CONCLUSION

As the proposed system is the automated human emotion recognition system which successfully displays emotion from an image uploaded by a user by matching it with our trained data set. In case when our trained datasets is matched with the uploaded image our system will show the output otherwise not. The input to the system is an image of a human emotion or a video from our database. The output is the emotion of the uploaded image or the frame taken from the video. Although facial expressions often occur during conversations, none of the cited approaches did consider this possibility. In the proposed approach emotion is depicted from an image as well as from video and target data is matched with source data in order to check or results over the trained data set. To check over the performance regression plot is another criterion. Confusion matrix is another measure to verify performance measure which shows percentage of right classification with respect to percentage of wrong classification.

. It is difficult to speculate on the relative usefulness of untested ideas, but it is worthwhile discussing potentially new directions that this research could take none the less. Despite the success of the research in the thesis, there is still much room for improvement for the adaptation and consistency modeling techniques presented in this paper. Three areas where future research could be directed are:

1. There are a number of applications where human emotion can be explored driving and monitoring.
2. Studying human psychology.
3. We have detected emotion from video on a frame work can be carried out to get emotion on each frame.
4. Emotion can also be detected from side images of human face.

REFERENCES

1. Tong Zang, WenMing Zeng "A Deep Neural Learning Network Driven feature Learning Method for Multiview Facial Emotion Recognition Method", Transactions on Multimedia, Dec 2016.
2. T. Ahonen, A. Hadid and M. Pietikainen, "Face description with Local Binary Patterns", Application to Face Recognition. Machine Vision Group, University of Oulu, Finland, 2014.
3. R. Gottumukkal and V.K. Asari, "An Improved Face Recognition Technique Based on Modular PCA Approach" Pattern Recognition Letters, vol. 25, pp. 429- 436, Mar. 2013.
4. M. Valstar, B. Jiang, M. Mehu, M. Pantic, and S. Klaus, "The first facial expression recognition and analysis challenge," in Automatic Face and Gesture Recognition, 2011.
5. A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in Proceedings of the ACM International Conference on Image and Video Retrieval, 2011.
6. V. Ojansivu and J. Heikkil, "Blur insensitive texture classification using local phase quantization," in Image and Signal Processing, ser. Lecture Notes in Computer Science, 2010.
7. T. Ahonen, A. Hadid, M. Pietikainen and T. M aenpaa. "Face recognition based on the appearance of local regions", In Proceedings of the 17th International Conference on Pattern Recognition, 2010.