IJITKM Volume 7 • Number 1 • December 2013 pp. 128-131 (ISSN 0973-4414)

Facial Expression Recognition Using Appearance Approach

Shalu Gupta, Sonit Singh
ECE Department, Lovely Professional University, Jalandhar, Punjab(India)
shalugupta54@gmail.com, sonit.15840@lpu.co.in

ABSTRACT
This paper provides a new approach to recognize facial expressions. In this paper, facial expression recognition is based on appearance based features or we can say that low level features. We used two different approaches to categories the expression into seven different classes. These classifications based on Scale Invariant Feature Transform (SIFT) and Local Gabor Binary Filter (LGBP). First, SIFT feature vector are extracted. We also perform the LGBP to extract the low level features. LGBP used 5 frequencies and 8 orientations. Our analysis based on JAFFE database. The experimental results show that the average recognition rates of different expressions, which is 88.4%.
Keywords- Facial expression recognition; SIFT; LGBP.

1. INTRODUCTION

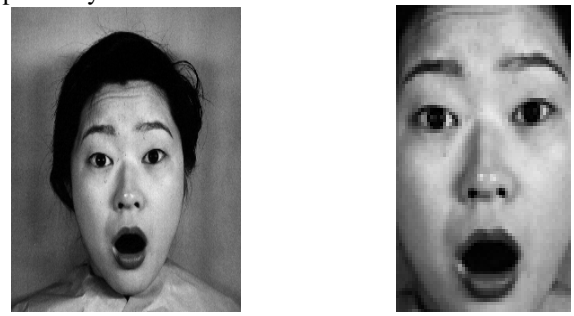
Facial expression recognition has become very crucial task in the field of robotics and computer visions. Robotic field has grown quickly in recent years. The applications of facial expression recognition are increasing day by day. The different areas which is covered by theses applications such as HMI (human machine interaction),computer vision design. Machine is more related to robots such as industrial robots, protection robots. With the help of expression robot can interact with their user in the more familiar way. Communication can be split into two parts. First, verbal communication and second is non verbal communication. An expression comes under the category of non verbal [1] communication. As we know one or two year's child can't speak but still he or she can communicate. So facial expression becomes an interesting field for research. There are two approaches to recognize expressions. These approaches are classified on the basis of types of features which are used in experiment. First, includes geometric feature which is oldest method. Active Appearance Method and Active Shape Model are the examples of geometric based features methods. Second method includes appearance [2] based features. SIFT LGBP, Linear discriminate analyses (LDA) [3], Principle Component Analysis (PCA) are the example of appearance based features methods. Both approaches have their own advantages and disadvantages. The biggest disadvantage of geometric [4] based features is reduction in spatial information. Another problem is high degree of ensemble rate. If matching of the models is not accurate then the overall performance degrades. G. Dorko and C. Schmid [5] provides spatial face ratio.
The low level features like texture [6] represents characteristics of facial image. In this paper, we propose a facial expression recognition method based on person – specific approach. The scale invariant feature transforms (SIFT) Descriptors extracted on dense facial points. 128 dimensional features vector is used in this method. Local Gabor Binary Pattern (LGBP) is 40 Gabor filter. Gabor wavelet provides better representation of neuron and neural field. Local Binary Pattern (LBP) has advantage of illumination invariant [7].

2. DATABASE

Facial expression recognition has been performed on JAFFE [8] database. This database is available publicly and anyone can access it. There are seven basic expressions (neutral (NA), happiness (HA), sad(SA), fear (FE), disgust (DI), anger(AN), surprise(SU))presented in the JAFFE database. The JAFFE database has total 213 images of seven different expressions.

3. FACE PREPROCESSING

Face preprocessing or image preprocessing includes two basic operations. First, it gives background information and alignment of face image. Second, it includes local information of an image. We obtain the images of uniform shape and size.fig 1 shows the surprise expression of original and localized image respectively of JAFFE database.



a .Original image b. Localized image
Fig 1. Pre-processing of face image

4. FRAME WORK

4.1 Feature Extraction

In the initial step we resize the images of JAFFE database. In this work we use resolution of 128 x 128 for every image. Although the original resolution of each image in database is 256 x 256. Then we extract the features of the images. Our feature vector consist of two features such as SIFT and LGBP features. A brief description of above two algorithms is given below.

4.1.1 SIFT has advantageous in case of variations in the image scale, rotation and illumination. In case of SIFT a Gaussian histogram is obtained. This histogram is SIFT descriptor. Now we use the feature vector which is available after reshaping the Gaussian histogram [9].Facial expression recognition using SIFT has given in [10]. SIFT has some problem Lowe [11] shows some 3D projection of images. It states SIFT has some classification problem with standard learning algorithm like support vector machine, which has problem of margin selection. In this work we use 8 orientations and total features in the vector have 6784 points or elements. In order to enhance the performance on we use discriminate features of SIFT. In this matching is calculated on the behalf of correlation. If there is high correlation between two features of SIFT then we can say that matching has been done. Following are the some training steps. Training algorithm

1) Divide each facial expression image into array of 4 x 4 patch or 16 grids of features. 2) Now obtained the SIFT features for both training images and test images in database. 3) Calculate the discriminative values for entire features and assign the preference to the SIFT features for images in the grid. 4) Place only that features which has some significant discriminative values or rank. 5) Repeat step 2 and 3 for all the images in the database.

4.1.2 LGBP consists of two approaches. First, Gabor filter and second, Local binary pattern. Both of these two approaches are explained briefly in the following sections. 4.1.2.1 Gabor filter gives us low level features. Gabor image is obtained by the convolution of original image and Gabor filter. In this work we use 40 Gabor filters. Gabor filter is derived by the following equation.

$$F_k(z)=k^2/\sigma^2\ e(\ -k^2z^2/2\sigma^2)(\ e^{ikz}-e^{-\sigma 2/2}) \qquad (1)$$

Where z(x,y) is spatial location of each pixel in Gabor filter. $F_k(z)$ is gabor kernel. It has two main parameters such as scale and orientation [12]. In this work we can define $k_v = (\frac{\Pi}{2}, \frac{\Pi}{4}, \frac{\Pi}{8}, \frac{\Pi}{16}, \frac{\Pi}{32})$ and $\phi_u = k\ \Pi/\ 8$ where k = {0,1,2,3,4,5,6,7}.which constitute 40 Gabor magnitude. The last term which is $e^{(\ \sigma^2/2\ )}$ used for compensation of d.c value. It is caused by nonzero [13] value of cosine value. Now input image and convolution of gabor kernel is obtained. Suppose input image is G(Z) have multiple scale [14] . Now the convolved image characterization is represented by $O_k(Z)$. If Z=(x,y) then O(x,y)

$$O_k(x,y) = F(x,y) * G(x,y) \qquad (2)$$

By using different values of orientations and frequencies we can obtain multi orientation and resolution analysis.

Above description is suitable for two dimensional Gabor filter for feature extraction.

4.1.2.2 LBP gives the local information of face space, or micro partition information. It requires less time to perform their operation. It computes their operation in one scale and reached to raw data which is required for further operation. So Local binary pattern plays a vital role in case of those images which have low resolution. It gives great result if input has low resolution. The local binary pattern has been calculated the local factor of an image with the help of non parametric descriptors. Local binary pattern convert a facial image into different small regions and histogram is calculated for individual blocks. Then each histogram is concatenated to entire histogram. The center value of the sub block is subtracted from each value of the block and counted in clockwise direction. The operator labels the pixels of an image by thresh holding the neighborhood of each pixel with the center value and considering the result as a binary number. LBP of pixel q of value i_q with neighbor $\{i_k\ k=0,1,2,3,4,5,6,7\}$

$$\sum_{k=0}^{7} \delta(fk - iq)\ 2^k \qquad (3)$$

Where $c(x) = \{1, \text{if } x \geq 0; \text{otherwise } 0\}$

If centre pixels is greater than pixels is encoded as one else encoded as zero. The very first function of local binary pattern to assign a special code to each every pixel, these pixels are called local binary codes (LBC).

28	100	140
50	110	60
90	150	80

Threshold

0	0	1
0		0
0	1	0

Fig 2 Example of LBP operator

Here threshold point is 110. Its binary equivalent is 00100100. In this case we have 4 transitions. If there are 2 transitions then it called a uniform LBP.

4.2 Classifier fusion

Different facial expression has been classified into their respective classes. Our basic approach for classification is K-nearest neighbor. Mean Distance between [15] the two feature vector of training and test images are calculated. First, Face pre-processing has been performed for both images than low level features (SIFT + LGBP) are calculated. One method for classification can be minimum pair distance but there is

problem of ambiguity. So in our approach we use KNN [16].

The output of classifier gives seven different expressions.

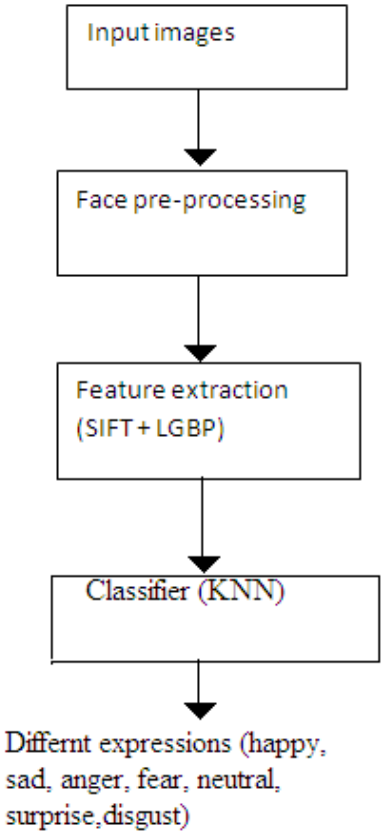


Fig. 3 Block diagram of proposed approach

In the above diagram feature extraction block shows our main approach which leads to SHIFT + LGBP. It states that SHIFT features are extracted from original images after this Local feature of gabor magnitude is obtained, which has the qualities of both of these features. When two or more feature vectors are used then performance changes their scale.

5 EXPERIMENTAL RESULTS

In our work we perform expression recognition in two steps. First, we use LGBP only for feature extraction. Table 1 shows the performance of LGBP on the JAFFE database. We perform our operation on 210 total images where we use different images of different expressions for the training purpose. Fifteen images of each expression for testing purpose. Table 1 confusion matrix for recognition based on LGBP features With the help of LGBP features performance is excellent in case of neutral expression but it is worst in case of fear expression. In case of fear expression it confused with disgust expression.

LGBP		Predicted expressions %						
		HA	SA	AN	DI	NU	SU	FE
Actual expressions	HA	80	0.00	0.00	20	0.00	0.00	0.00
	SA	0.00	73.33	0.00	0.00	13.33	0.00	13.33
	AN	0.00	13.33	86.66	0.00	0.00	0.00	0.00
	DI	6.66	0.00	0.00	66.66	13.33	0.00	13.33
	NU	0.00	0.00	0.00	6.66	93.33	0.00	0.00
	SU	6.66	0.00	0.00	6.67	6.67	80	0.00
	FE	0.00	13.33	0.00	20	0.00	6.66	60

At second step we use the features of both SIFT and LGBP. Now table 2 shows the performance of SIFT + LGBP on the same JAFFE database.

Table 2 confusion matrix for recognition based on SIFT and LGBP features

SIFT + LGBP		Predicted expressions %						
		HA	SA	AN	DI	NU	SU	FE
Actual expressions	HA	80	0.00	0.00	13.33	0.00	6.66	0.00
	SA	0.00	86.66	0.00	0.00	13.33	0.00	0.00
	AN	0.00	6.66	86.66	0.00	6.66	0.00	0.00
	DI	6.66	0.00	0.00	93.33	0.00	0.00	0.00
	NU	0.00	13.33	0.00	13.33	73.33	0.00	0.00
	SU	0.00	0.00	6.33	0.00	0.00	93.33	0.00
	FE	0.00	13.33	0.00	6.66	0.00	6.66	73.3

When features are extracted with SIFT and LGBP performance has excellent result for two expressions such as disgust and surprise. The recognition rate for fear is also increased. So it is clear from the above results combination of two features gives improved results.

6. CONCLUSION

Our proposed method is better than previous LBP and SIFT alone. This approach signifies in such situations where variation in illumination and scale is present. Our experiment based on person specific images. If the effect of rotation is present in the images then this approach gives satisfactory results. The small numbers of training images for each expression in the database are provided even though recognition rate is good. The computational complexity is same as previous work. If we use principle component analysis (PCA) or independent component analysis (ICA) computation time can be reduced. Our proposed method is limited to pre processing. If more than above two features are used may give more superior results.

REFERENCES:-

[1] A. Mehrabian and Wiener, "Decoding of inconsistenty communication," j. Per Soc. Psychol., vol.6, no. 1, pp. 109-114, May 1967.

[2] Z. Wen and T. Huange, "Capturing sutable facial motions in 3D face tracking," in proc. IEEE ICCV, 2003, pp. 1343-1350.

[3] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 711–720.

[4] T. seneschal, V. Rapp, H. Salam, R. Seguier, K. Baily, L. Prevost, "Facial action recognition combination heterogeneous feature via multikernal learning," IEEE Transaction on system, man, and cybernetics-part B, vol. 42, No. 4, August 2012

[5] G. Dorko and C. Schmid, "Object class recognition using discriminative local features". IEEE PAMI, 2004.

[6] M. F. Valstar, M. Mehu, B. Jiang, M. pantic K. schrer, "Meta – Analysis of first facial expression recognition challenge "IEEE Transaction on system, man, and cybernetics-part B, vol. 42, No. 4, August 2012.

[7] C. Shan, S. Gong , Peter W. McOwan," Robust Facial expression recognition using local binary pattern",IEEE 2005.

[8] Vaibhavkumar, J. Mistry, Mahesh , M. Goyani, "A literature survey on Facial Expression Recognition using Global Features," International Journal of Engineering and Advanced Technology ISSN: 2249 – 8958, Volume-2, Issue-4, April 2013.

[9] A. Majumdar and R. K. Ward, "Discriminative scale features for face recognition," 978-1-4244 IEEE 2009.

[10] U. Tariq and T. S. Huang , "Features and fusion for expression recognition - A comparative analysis," 978-1-4673 IEEE 2012.

[11] D. Lowe, " Distinctive image features from scale invariant key points".Internaional Journal of Computer Vision, 60(2):91–110, 2004.

[12] J. Whitehill, C. Omlin, "Haar features for facial action coding system action unit recognition". In Proc. IEEE Intl Conf. Face and Gesture Recognition, 2006.

[13] S. Fazli, R. Afrouzian, H. Seyedarabi ," High performance facial expression recognition using gabor filter and probablisitc neural network, " 978-1-4244-4738 IEEE 2009

[14] F. Y. Shih, C. Chuang, "Automatic extraction of head and face boundaries and facial features", Information Sciences, Vol. 158, 2004, pp.117-130.

[15] T. W. Marian, S. B. Javier, R. Movellan ," Facial Expression Recognition Using Gabor Motion Energy Filters" 978-1-4244-7030 IEEE 2010.

[16] J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", IEEE Transaction. Acoustic, speech and signal processing, Vol. 36, 1988, pp.1169-1179.