Survey of Various Feature Extraction and Classification Techniques for Facial Expression Recognition

AYYAZ HUSSAIN[1], MUHAMMAD SHAHID KHAN[1], MUHAMMAD NAZIR[2], M AMJAD IQBAL[3]
[1]Faculty of Basic and Applied Sciences, Department of Computer Science & Software Engineering, Islamic International University Islamabad, PAKISTAN
[2]National University of Computer & Emerging Sciences FAST_NU Islamabad, PAKISTAN
[3]Faculty of Information Technology, University of Central Punjab Lahore, PAKISTAN

ayyaz.hussain@iiu.edu.pk, mrkhan.inbox@gmail.com, muhammad.nazir@nu.edu.pk, amjad.iqbal@ucp.edu.pk

Abstract: - The purpose of this paper is to present a comprehensive study of latest and most famous facial features extraction techniques and as well as classification techniques. We studied these techniques in two different perspectives; one is spatial domain and other is frequency domain. We found many advantages and disadvantages of each technique inside one domain and as well in between different domains. We observed that Local Binary Pattern is a new technique and now it is becoming very famous technique in spatial domain. LBP simply knows about micro patterns using the comparison with neighbour pixel grey scale values. Lot of work on LBP has yielded its different extensions which have optimized the base concept of LBP operator. Frequency domain covers the techniques which transform images into frequency domain and use either cosine or sine waves to extract the facial features. This method is a very strong and precise that only two to three features have the ability to describe the facial expressions. We have also studied different classification techniques in domain of facial expressions classification. In most of the solutions classification is done through KNN classifier. K Nearest Neighbour is a successful and non-parametric technique of machine learning in supervised learning techniques.

Key-Words: - Facial Features, Local Binary Pattern, Local Directional Pattern, Classification, Facial Expression Recognition, K-Nearest Neighbor

1 Introduction

Automated process of facial features extraction forwards with the steps of image normalization, noise removal and finally facial features extraction. There have been a lot of research work done in this area and still number of researches going on to optimize and to increase efficiency of existing techniques which has resulted many new techniques but still it is very challenging and complex research area due to variant conditions of facial images like presence of large number of variations in facial expressions itself, facial details, face poses and illumination conditions. Extraction of precise, useful and important facial features are very helpful in increasing the efficiency of a classifier by decreasing the classifier complexity where as collection of unimportant and poorly extracted features leads to poor classification results and create a scenario where a very good classifier cannot even perform efficient classification. A number of machine learning techniques are available to perform classification efficiently in different fields but these methods trade off for complexity, performance, and time-space consumption due to dataset variations, dataset size and dataset density.

Human faces have too much diversity due to varying facial details from one person to other person, and from one region people to another region people which makes classification a very challenging task. Human facial features are broadly categorized into two areas [6]; first are geometric-based features: shapes and locations of facial micro organs like nose, eyebrows, eyes and lips and second are surface-based features: face appearances changing and skin textures like face wrinkles and furrows. Different methods are present for features extraction and we will discuss these methods here.

2 Literature Review

2.1 Local Binary Pattern

Local Binary Pattern (LBP) descriptor [1] is an efficient facial features extraction method; it extracts information from neighbouring pixel values and develops histogram of the image. This is a non-parametric operator and describes the local spatial structure of an image. It calculates a bit-code from binary derivates of pixels and then finds the difference of central pixel with its neighbouring pixels, arrange these differences in an ordered form, and finally this bits pattern is converted into decimal value which is the new LBP code for the central pixel. Basic LBP operator works for 3×3 pixels. It is described by the following diagram;

129	125	199
120	110	90
139	119	87

→

1	1	1
1		0
1	1	0

Binary Code: 11100111

LBP Code: 231

Fig. 1 Basic LBP Operator

Above figure shows grey scale values of 3×3 pixels and LBP code is calculated using the following formula.

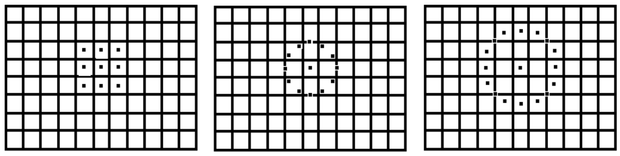
$$LBP(x_c, y_c) = \sum_{n=0}^{7} S(i_n - i_c)2^n$$

$$S(x) = \begin{cases} 1 \; if \; (x \geq 0) \\ 0 \; if \; (x < 0) \end{cases} \tag{1}$$

Here x_c *and* y_c shows the position of centre pixel, i_n and i_c are gray scale values of surrounding pixels and central pixel respectively. After labelling the image with LBP codes, image histogram is generated which helps to recognize micro patterns in image like eyes, nose and lips etc.

2.2 Local Binary Pattern Extension

One of the LBP extensions [17] is very efficient with the concept of circular neighbourhood with varying radius and with varying number of neighbourhood pixels.



(P=8, R= 1.0) (P=12, R= 1.5) (P=16, R=2.0)

Fig. 2 LBP Operator Extension [17]

This figure shows radius R=1.0, R=1.5, R=2.0 and pixels count like P=8, P=12, P=16. This form can be represented by the following formula;

$$LBP_{P,R}^{u2} = \sum_{j=0}^{p-1} S(g_j - g_c)2^j$$

$$S(x) = \begin{cases} 1 \; if \; (x \geq 0) \\ 0 \; if \; (x < 0) \end{cases} \tag{2}$$

Here g_j and g_c represents the neighbourhood pixels and central pixel respectively.

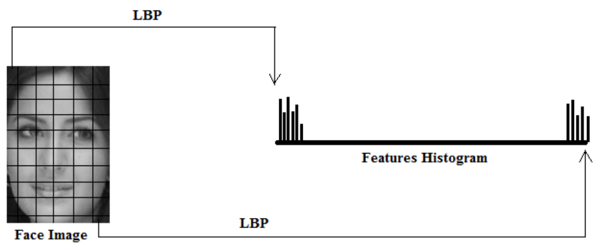
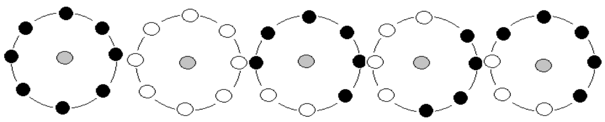


Fig. 3 Small regions of face are used to extract histograms which are then concatenated into a single histogram to show complete face feature histogram.

LBP codes for small regions are calculated to construct histograms which are combined to show global histogram. The circular results show different micro patterns of facial image like spot, line end and edge etc.



Spot Flat Line End Edge Corner

Fig. 4 LBP detectable texture primitives

Black circles show zeros and white circles show ones.

2.3 Local Direction Pattern (LDP)

LDP [3] uses edge responses of each pixel in eight directions, and generates LDP code. Kirsch masks are used in eight directions ($M_0 \sim M_7$) with the central pixel position as shown in Fig. 5.

$$M_0 = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad M_1 = \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad M_2 = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad M_3 = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad M_5 = \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad M_6 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad M_7 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

Fig. 5 Kirsch edge response masks in eight directions

Each direction mask M shows edge response of pixel in that particular direction but every direction response is not equally important because only edge or corner which shows high response are important for calculating LDP code. This algorithm finds out K prominent responding directions and set value as 1 while remaining 8-k directions are set as 0. An example is described in the following figure.

85	32	26
53	50	10
60	38	45

Index	m_7	m_6	m_5	m_4	m_3	m_2	m_1	m_0
Value	161	97	161	537	313	97	-503	-393
Rank	6	7	5	1	4	8	2	3
Code Bit	0	0	0	1	0	0	1	1
LPD Code	19							

Fig. 6 LDP Code Generation

Mask values are derived after applying Kirsch masks and are placed on masks index $m_0 \sim m_7$. If we carefully observe then value (537) has high variance and then (-503), and (-393) are showing high responses so this method assigns these values as 1 and remaining as 0 so the LDP code is 00010011 which means 163. In this way complete image is processed for LDP code and finally face micro patterns are described by LDP histograms as shown below.

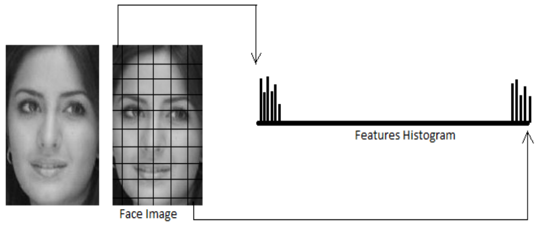


Fig. 7 Facial representation using LDP histogram

2.4 Local Ternary Pattern

LTP is also an extension of LBP operator and the difference is in basic LBP operator it has 59 bins of histogram while in LTP bins are further reduced to 24 only which give very good results. As in LBP, it also uses the concept of uniform patterns; which allows maximum two bit wise transitions in code. It uses the following formula to calculate LTP code.

$$LTP(x_c, y_c) = \sum_{n=0}^{7} S(g_n - g_c) 2^n$$

$$S(x) = \begin{cases} 0 \; if \; (x < -\theta) \\ 1 \; if \; (-\theta \leq x \leq \theta) \\ 2 \; if \; (x > \theta) \end{cases} \tag{3}$$

Here the central pixel is evaluated against the 8 neighbours using a threshold θ. LTP sows grey scale values invariance because, similar to LBP, it actually uses the difference between grey scale values rather than original grey scale values. Let's take an example to calculate LTP.

129	125	199
120	110	90
139	119	87

2	2	2
2		1
2	2	0

LTP:
22102222

Fig. 9 Calculation of LTP Code

2.5 3D Local Binary Pattern

This method [16] uses a different approach in comparison with 2D where a central pixel is checked against its circular neighbours and in 3D the neighbours of central pixel are defined in a sphere.

2.6 Discrete Cosine Transform

Discrete Cosine Transform (DCT) uses sine functions to transform images into frequency domain. It is a famous technique to reduce dimensionality and used for image compression generally. DCT method selects only sine waves from transformed image in frequency domain. The following equation is used to calculate DCT coefficients *C (u, v)* of image *f (x, y)*.

$$C(u,v)=\frac{2}{\sqrt{mn}}(\alpha(u)\alpha(v))\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}f(x,y)\cos\left[\frac{(2x+1)u\pi}{2m}\right]\cos\left[\frac{(2y+1)v\pi}{2n}\right]$$

Discrete Cosine Transform Coefficient Formula
DCT has very good quality to pack high information which makes it very robust for image compressions. DCT arranges its coefficients in priority order and high significant coefficients are chosen first which show high variance, this variance decreases with high frequency, so for dimension reduction, we usually select low frequency coefficients.

2.6 Machine Learning Techniques:

In Local Binary Pattern [2], it uses KNN technique to perform classification using Euclidean distance from X to Y. Let us suppose that it is m-dimensional LBP feature vector and the Euclidean distance will be as following;

$$d(X,Y)=\sqrt{\sum_{i=1}^{m}(x_i-y_i)^2}$$

In Local Directional Pattern [4], it uses two methods of machine learning techniques; Template Matching and Support Vector Machine (SVM). In Template Matching, during training stage a template for each class is generated using its average histogram. It then compares test histogram with template histogram for classification. It has used weighted chi-square method for similarity measure between test & template histogram. The purpose of weighted version is to prioritize some face regions with high score if those regions contribute more towards expression classification like eyes & lips got high weights. Chi-square method is as below;

$$x_w^2\left(SLH^1,SLH^2\right)=\sum_{i,j}w_i\frac{\left(SLH_{i,j}^1-SLH_{i,j}^2\right)^2}{\left(SLH_{i,j}^1+SLH_{i,j}^2\right)}$$

Here SLH^1 and SLH^2 are histograms of template expression and histogram of test expression respectively, i represents region number, j represents bin number of i region and w_i is the associated weight of that region i.

In Local Directional Pattern [4], the second method used of machine learning is Support Vector Machine. It is a supervised learning technique which do mapping of data in a higher dimensional feature space. It then finds a linear hyper plane with maximal margin. This technique is very efficient for binary classification but multi-class classification can also be achieved by adopting the one-against-one or one-against-all techniques.

In Local Ternary Pattern, a non-parametric method known as G statistics (by Sokal & Rohlf) is used. A non-parametric is advantageous because there is no need to assume about feature distribution. Its equation is as below;

$$G(S,M)=2\sum_{b=1}^{B}S_b\log\frac{S_b}{M_b}=s\sum_{b=1}^{B}[S_b\log S_b-S_b\log M_b]$$

Here S is sample distribution & M is model distribution. S_b & M_b are probabilities of bin b in sample and model distribution respectively. B represents the number of bins in distributions. It simply uses nearest neighbour rule to classify unknown sample in model distribution.

3 Conclusion

Local Binary Pattern uses KNN with Euclidean Distance Method and is Rotation Variant with 256 histogram bins, 8 neighbours and is crisp operator, very powerful and gray-scale invariant. Local Binary Pattern Extension uses arbitrary circular neighbourhoods and is rotation Invariant and use only uniform patterns with large neighbourhood radii results in sparse sampling. Concept of uniform patterns reduces histogram bins up to 59. Local Directional Pattern uses Template Matching & Support Vector Machine for classification but Better in white noise by Gaussian Noise. Local Ternary Pattern uses G Statistics method for classification and Rotation Invariant and is Histogram Bins are 24 with 8 neighbours.

References:

[1] W. Wang, F. Chang, J. Zhao and Z. Chen, "Automatic Facial Expression Recognition using Local Binary Pattern," Proceedings of the 8th World Congress on Intelligent Control and Automation, Jinan, China, (2010)

[2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Recognition with Local Binary Pattern," Springer-Verlag Berlin Heidelberg, ECCV, LNCS 3021, pp. 469–481, (2004)

[3] T. Jabid, M. H. Kabir, O. Chae, "Facial Expression Recognition using Local Directional Pattern (LDP)," Proceedings of 2010 IEEE 17th International Conference on Image Processing, Hong Kong, (2010)

[4] S. M. Lajevardi, Z. M. Hussain, "Automatic Facial Expression Recognition: Feature Extraction and Selection," Springer-Verlag London Limited, (2010)

[5] A. Mehrabian, "Communication without words," Psychol. Today 2(4), 53–56, (1968)

[6] Tian, Y.L. Kanade, T. Cohn, J.F, "Facial Expression Analysis. Hand Book of Face Recognition," pp. 247–275 (2005)

[7] T. Yang, V. Kecman, "Face recognition with Adaptive Local Hyperplane Algorithm," Pattern Anal Applic,13:79–83, Springer-Verlag London Limited, (2010)

[8] P. Vincent, Y. Bengio, "K-local Hyperplane and Convex Distance Nearest Neighbor Algorithms," Advances in neural information processing systems (NIPS), vol 14. MIT Press, Cambridge, pp 985–992 (2002)

[9] L. SZ, L. J., "Face Recognition using the Nearest Feature Line Method," IEEE Trans Neural Netw, 10:439–443 (1999)

[10] W. Zheng, L. Zhao, C. Zou, "Locally Nearest Neighbor Classifiers for Pattern Classification," Pattern Recognition, 37:1307–1309 (2004)

[11] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification," John Wiley & Sons, Inc., New York, (2001)

[12] C. Shan, S. Gong, P. W. Mcowan, "Facial Expression Recognition based on Local Binary Patterns, A Comprehensive Study," Image and Vision Computing 27 803–816, Elsevier B.V, (2009)

[13] T. Ojala, M. Pietikäinen, T. Mäenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (7) 971–987 (2002)

[14] J. Shlens, "A Tutorial on Principal Component Analysis," (2005)

[15] K. Cabeen, P. Gent, "Image Compression and the Discrete Cosine Transform," College of the Red Woods, (2000)

[16] H. C. Akakin, A. A. Salah, L. Akarun, B Sankur, "2D/ 3D Facial Feature Extraction," SPIE-IS&T, Vol. 6064 60641D-1, (2006)

[17] T. Ojala, M. Pietikäinen, T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) 971–987, (2002)