Local Polynomial Approximation-Local Binary Pattern (LPA-LBP) based Face Classification

Rakesh Mehta*[a], Jirui Yuan*[a], Karen Egiazarian*[a]
[a]Department of Signal Processing, Tampere University of Technology, Tampere, Finland, FI 33120

ABSTRACT

In literature of face recognition many methods have been proposed which extract features at multiple scales for robust classification. In this paper, we proposed a novel method which utilizes Local Polynomial Approximation (LPA) techniques to capture the directional information of the face image at different scales. LPA based filters are used to obtain directional faces from the normalized face images at multiple scales. Since face image is spatially varied and classification works better when local descriptors are used, we incorporate Local Binary Pattern (LBP) operator to obtain LPA-LBP maps. Blockwise processing is done on LPA-LBP maps to capture the local regional relation among the pixels. Then, finally, Support Vector Machine (SVM) classifier is learned in LPA-LBP feature space for face classification. The final descriptor contains information extracted from different levels and, thus, results in high classification accuracy of the faces. Experiments done on Yale and ORL datasets demonstrate that the proposed method has higher classification accuracy than previously proposed methods.

Keywords: LPA, Face Classification, Face Representation, LBP, Multi-scales method

1. INTRODUCTION

FACE classification has been widely studied in recent years. Its practical application in security and video-surveillance has promoted the studies in this direction. Given a set of training images and their corresponding classes, the face classification algorithm should be able to classify the test image into a specific class. An important issue in this process is to find an efficient method for face representation robust to illumination variation and pose changes. Methods based on holistic approach such as PCA and LDA have been studied widely but lately local descriptors have gained attention due to their robustness to illumination and pose changes. In this paper we propose a face representation method based on Local Polynomial Approximation (LPA) and Local Binary Pattern (LBP).

Principle Component Analysis (PCA) based method[1] projects an image into a space where the individual components are ranked according to their variance. Thus, a face is represented as a linear combination of most expressive faces. The shortcoming of this is that it is highly sensitive to illumination and pose changes. To overcome the illumination effect Fisherface[2] based recognition was proposed. This method utilizes the class based information of labeled face images to obtain a feature space which is robust to illumination changes. Although this method works better than PCA based methods on most occasions, it was shown[3] that the efficiency of Fisherface based method reduces significantly as the number of training samples are reduced. Because of the limitations of holistic representations, methods based on local descriptors gained wider interest. LBP operator[4] being a powerful tool for texture analysis found application to face classification also. For this purpose, the LBP operator is applied by segmenting an image into a fixed number of blocks so that a descriptor has local region based information. Because of its computational simplicity and robustness to illumination changes, LBP has been widely combined with many other methods. Another type of local descriptors widely used in face classification is a Gabor based descriptor. Gabor Filters are bandpass filters which are tuned to find the changes in the texture in an image. These filters enhance the edge contours, thus when applied to face images they result in enhancing eye, mouth, nose edge which are suppose to be the main regions of a face. Gabor feature were combined with LBP operator to extract the texture information at multiple scales from a face image to finally obtain LGBPHS[5] (Local Gabor Binary Pattern Histogram Sequence). But in this method each image corresponding to a specific orientation and scale was treated independently from the images of other scales and orientations during the computation of the final descriptor due to which the relation between different scales and orientation was not well captured. This problem was overcome by Lei's method[6] which uses the volume based LBP between neighboring images to completely capture the relation between different scales and orientations. Gabor based methods are however unable to capture the

Multimedia on Mobile Devices 2011; and Multimedia Content Access: Algorithms and Systems V, edited by David Akopian, Reiner Creutzburg, Cees G. M. Snoek, Nicu Sebe, Lyndon Kennedy, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7881, 78810N · © 2011 SPIE-IS&T · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.878796

directional information of the face. Heat kernel based face representation was proposed for classification purpose by Li and Hu[7]. Heat kernels were used at different time scales to capture intrinsic structural information of the face appearance. This method involves graph computation for face image and then performing matrix operation on it to get the final feature vector. For the image of size $m \times n$ the graph is of dimension $(m \times n) \times (m \times n)$ and performing operation on matrix of this size adds a computational complexity to it.

In this paper, we propose a novel method for face recognition. The outline of the method is as follows. The method is based on obtaining the approximations of face image using LPA filters for specific number of directions and scales. Filtering the face image with these filters, we can extract directional information of the face image at different scales. After extracting the directional information, LBP operator is applied on these set of images to obtain LPA-LBP map which completely capture the texture information from them. The LPA-LBP map is a holistic representation of the face. In order to have a local descriptor, LPA-LBP maps are divided into blocks and histogram is evaluated for each block. Then all the histograms are concatenated to form the final descriptor LPA-LBP Histogram Sequence (LPA-LBP-HS). Finally, an SVM classifier is learned in the LPA-LBP-HS space for the purpose of face classification.

2. FACE REPRESENTATION BASED ON LPA-LBP-HS

The face representation based on LPA-LBP-HS consists of following steps: 1) the LPA filters are computed for a fixed number of scales and the number of directions is set to 4 to keep the computational complexity low; 2) the normalized face images are filtered using these LPA filters; 3) LBP operator is applied on the filtered images to obtain LPA-LBP maps; 4) the LPA-LBP maps are divided into non-overlapping square blocks and histogram is computed for each block separately; 5) the histograms of the blocks of all the LPA-LBP maps are concatenated to form the final LPA-LBP-HS feature vector. All of these steps are discussed in details below.

2.1 LPA Kernels

A large amount of information in a face image is aligned in some specific directions. For example, face features like eyebrows (which are considered to be most important in face recognition), eyes and lips are aligned in horizontal direction, while face features such as eyeballs, nose and face contour are aligned in vertical direction. The idea here is to use directional filters from which we can extract this directional information of the face at different scales. Local Polynomial Approximation technique is utilized for generating these directional filters[8].

LPA is a powerful tool in signal and image processing. It has been utilized in a number of applications related to image processing such as denoising[9], image reconstruction[10], phase unwrapping[11], color filter array interpolation[12] etc. LPA is applied for linear filter design using a polynomial fit in a sliding window. It is based on nonparametric estimation of the signal. The signal can be represented as a linear combination of polynomials centered on the observation coordinates using the Taylor series expansion. Thus, the signal is represented as

$$y(x, X_s) = C_0 - C_1(x - X_s) + C_2(x - X_s)^2/2, \tag{1}$$

where X_s are coordinates of observations and constants C_0, C_1 and C_2 are unknown. Comparing the above equation with the Taylor series expansion of the signal, we get the estimates of the signal and its derivatives. For higher order derivatives we can use a higher order polynomial in (1). To formalize a local fit, a weighted least square criterion given below is used

$$J(x, C) = \sum_s w(x - X_c) e_s^2 \tag{2}$$

$$e_s = z_s - y(x, X_s) \tag{3}$$

In the above equation e_s indicates the error between the actual and the estimated value. Thus, we estimate the error over the neighborhood of the observation coordinates using a window function w(x) centered on the point of observation. Our aim is to minimize the squared error given in (2). Thus, we put the following constraint for obtaining the estimate of the signal and its derivatives

$$\hat{C}(x) = \arg\min_C J(x, C) \tag{4}$$

It can be observed from (4) that the coefficients C depend on x, thus the coefficients are different for different locations of x. The sliding window makes the coefficients varying on x. The estimates as well the derivative of the estimates can be given in the form of kernel operator

$$\hat{y}_h(x) = \sum_s g_h(x, X_s)\, z_s \tag{5}$$

$$g_h(x, X_s) = w_h(x - X_s)\phi_h^T(x - X_s)\psi_h^{-1}\phi(0) \tag{6}$$

$$\psi_h = \sum_s w_h(x - X_s)\phi_h(x - X_s)\phi_h^T(x - X_s) \tag{7}$$

Here ϕ is a vector of polynomial for LPA and the length of the vector is equal to the order of polynomial plus one. The kernels obtained from this method can be used to generate the estimates of the signal and its derivatives. We use the derivative estimate for generating the directional face. For a particular scale, if we have the kernel for one direction, then the kernels for other three directions can be obtained just by rotating it.

Figure1 shows the directional faces generated by applying LPA filters on the face image. Each row corresponds to a particular direction. It can be observed that in the first and third row of filtered images vertical features of the face are quite prominent because these were obtained by applying vertical filers. Similarly, the images in second and fourth row which were obtained by applying horizontal filters, have highlighted horizontal features of the face. Thus, using LPA filters for four different directions and different numbers of scales, we extract the complete directional information from the face image.

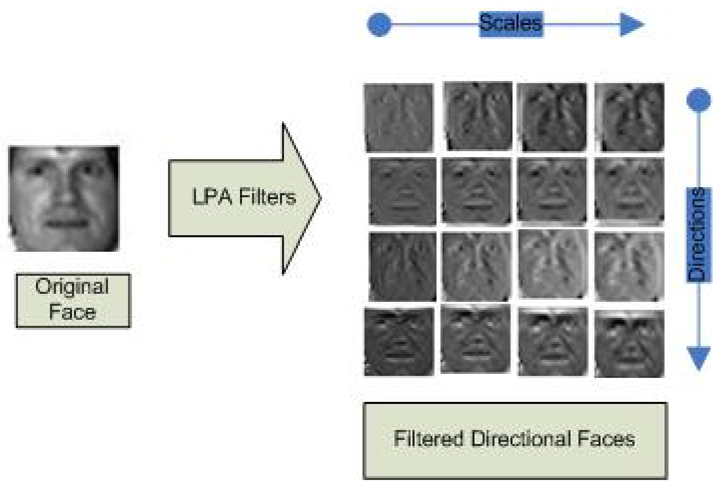

Figure 1. Face image is filtered using LPA directional filters with four directions and four scales.

2.2 LPA-LBP maps

Local Binary Pattern is a powerful tool for texture analysis of an image. In order to further extract the information from the LPA filtered images LBP is utilized. Many variants of LBP are proposed but we use the original LBP operator which labels the pixel of an image by thresholding the 3x3 neighborhood of each pixel f_p (p= 0, 1, …, 7) with the central value f_c and considering the number as a binary value

$$S(f_p - f_c) = \begin{cases} 1, & f_p \geq f_c \\ 0, & f_p < f_c \end{cases} \tag{8}$$

After thresholding the neighboring pixels with respect to the central pixel for each central pixel, the LBP pattern is obtained as

$$LBP = \sum_{p=0}^{7} S(f_p - f_c)2^p \tag{9}$$

LBP can be seen as the first order derivative of the image pixel in different directions. Number of directions can be considered as eight since the difference is taken between a pixel and its eight neighbors and the final descriptor which is obtained takes all these values into consideration after thresholding them. It is able to extract the information from the images very efficiently both in terms of speed and discrimination performance and has been used for face classification purpose. Thus we also use LBP to extract complete information from the directional faces. LBP map is obtained for all the filtered directional images and hereafter they are called LPA-LBP maps.

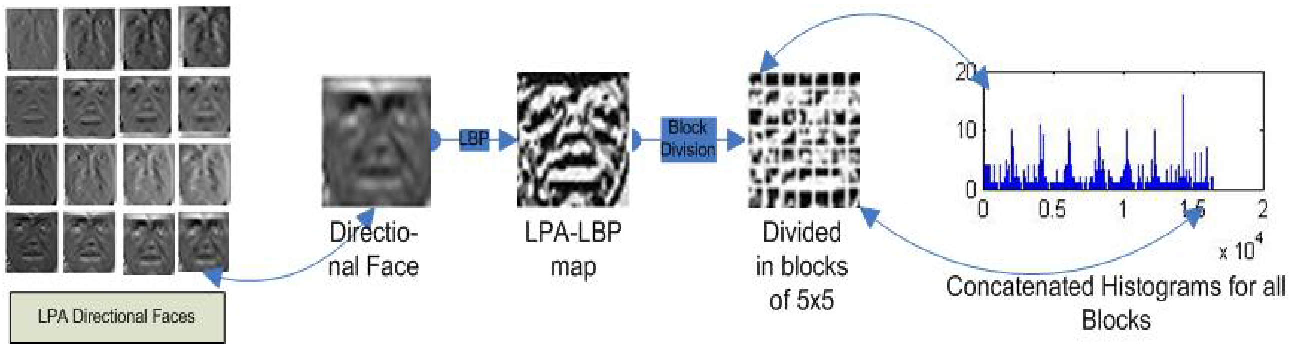

Figure 2. The figure shows how the directional faces are further processed. First, on each directional face we apply the LBP operator in order to obtain the LPA-LBP map. Then, LPA-LBP map is divided into blocks of size 5x5. Histogram is obtained for each block of the directional face. Finally, all the histograms are concatenated to form a histogram sequence for a single directional face.

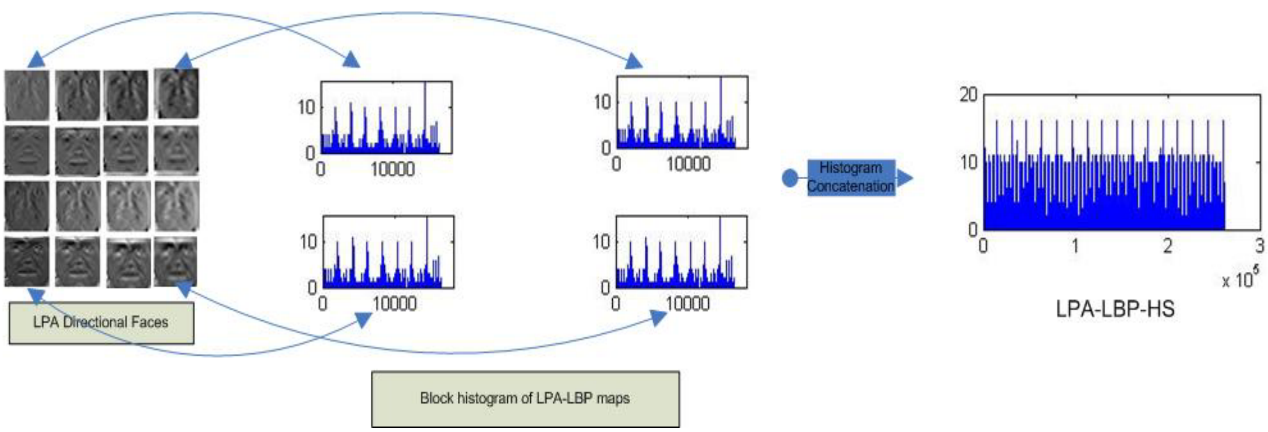

Figure 3. Concatenated histogram sequence is obtained for all the directional faces by the procedure shown in Figure 2. To form the final feature vector LPA-LBP-HS, we concatenate the histogram sequence of all directional faces.

2.3 LPA-LBP-HS

LPA-LBP maps give the holistic representation of the face. In order to achieve an efficient classification face should be represented as a localized feature vector where separate regions of face are encoded separately. Therefore we divide the LPA-LBP maps into non-overlapping blocks of fixed size and then compute the histogram of each block. Block size used for partitioning LPA-LBP map is also an important parameter and it should be properly set according to the size of the image. If the block size is too big then the blocks will not capture local information from the various directional features of the face image. The horizontal directional features such as eyes will get mixed up with vertical directional features such as nose. If the block size is too small then the relation between neighboring pixels will not captured and it will almost be equivalent to extraction of information from the pixel level which has been already done by using LBP. It is also important to remember that computational complexity increases considerably as the block size is reduced. The choice of a block size with respect to the size image is further discussed in the results section.

Thus, after fixing the block size, a histogram is computed for each non overlapping block for each LPA-LBP map. Histograms of all the blocks of a single LPA-LBP map are then concatenated to obtain a feature vector for a single LPA-

LBP map representing a specific direction and scale. Finally, the histograms obtained from different LPA-LBP maps are concatenated to form the final feature vector.
The histogram for a gray level image can be defined as

$$h_i = \sum_{x,y} I\{f(x,y) = i\}, i = 1,2, \dots L-1, \tag{10}$$

where the range of gray level is given as [0,L-1], *i* is the i^{th} gray level, h_i is the number of pixels with gray level i and

$$I\{A\} = \begin{cases} 1, & A \; is \; true \\ 0, & A \; is \; false \end{cases} \tag{11}$$

If we assume that each filtered image is divided into m regions and the histogram for a k^{th} block is represented as h_k, then the feature vector for a specific scale *s* and direction *d* is given as

$$H_{s,d} = \{h_1, h_2, \dots h_m\} \tag{12}$$

Figure 2 shows how for a single directional face we obtain feature vector $H_{s,d}$. The final feature vector LPA-LBP-Histogram Sequence (LPA-LBP-HS) is obtained by concatenating the feature vectors $H_{s,d}$ for all scales s = 1,2,3,4 and for all directions d = 1,2,3,4 as shown in Figure 3.

$$LPA - LBP - HS = \{H_{1,1}, \dots H_{1,4}, H_{2,1} \dots H_{4,4}\} \tag{13}$$

For example, if four directions and four scales are chosen, then sixteen LPA-LBP maps will be obtained. If each LPA-LBP maps is divided into 64 blocks, then the total number of blocks is 1024 for all 16 LPA-LBP maps. The histograms for all these 1024 blocks for a single face image are concatenated to form the final feature vector. The final feature vector which is used for face representation encodes the face information at three different levels. First, the global directional information extracted by the directional filters at multiple scales, secondly, the texture based information at the pixel level is extracted by the LBP operator, and finally, the regional information of the face is extracted by dividing the LPA-LBP maps into blocks of fixed size. This information encoding at multiple level helps in robust face representation of face image for classification purpose.
Face classification is performed using a Support Vector Machine (SVM) classifier. A multiclass SVM classifier with linear kernels is trained in LPA-LBP-HS space by using the LIBSVM tool. SVM classifier stores certain number of support vectors which lie near the margin and use these vectors for a prediction.

3. EXPERIMENTS

To check the performance of the LPA-LBP-HS in face classification, a several tests were performed on Yale and ORL datasets. Yale dataset consists of 165 images of 15 individual. There are 11 images of each individual, one per different facial expression or configuration: center-light, with glasses, happy, left-light, without glasses, normal, right-light, sad, sleepy, surprised, and wink. The Yale database images have a fixed pose but different light illumination condition and expression. ORL dataset consist of images from 40 different individuals with 10 images of each person. The images were normalized and cropped. All the face images were resized into images of 35x35 pixels for properly characterizing the directional information. Datasets were randomly partitioned into training dataset and testing datasets. SVM classifier with a linear kernel was trained using the training data. The process was repeated 50 times and the average accuracy of a classification was calculated. A number of factors were varied while testing which is discussed in details below. Initially the parameters of the LPA filters are set by doing tests on the Yale dataset. Then, using the parameter value obtained from these tests, the final tests are performed on both Yale and ORL dataset.

3.1 Number of Scales

The scale represents the size of the window which is also the size of the kernel used to filter the face image. The size of window is an important parameter for the local approximation. If it is too large, it prescribes equal weights to all residual and nonparametric estimate coincides with the parametric estimate. If the window is of minimal size then estimation is same as the point of observation. The number of scales of LPA filters varies from 2 to 5. As more scales are used, more directional information from the face images are extracted. But if the number of scales is too large then after a certain point the increase in the dimensionality due to a scale leads to a large number of features which further reduces the accuracy. The results for a classification on Yale database for the different scales are shown in Figure 4. The average accuracies for the scales were found out to be (93.15, 93.57, 94.60 and 93.27). It can be observed that the accuracy increases till a particular point after which it starts decreasing.

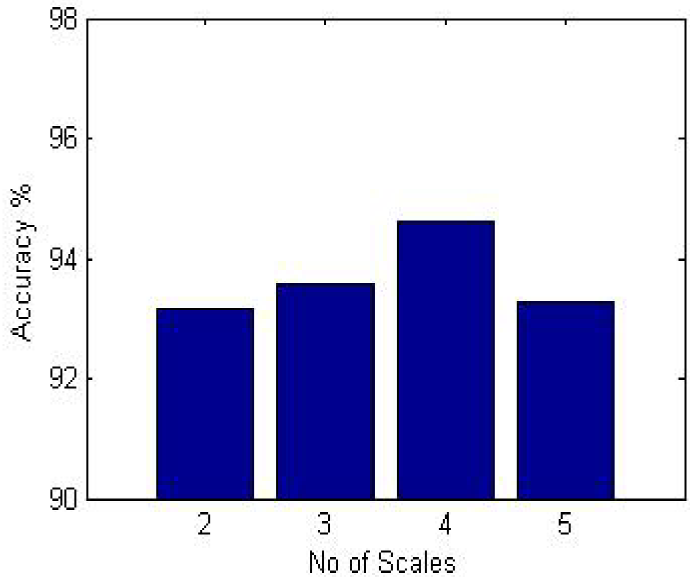

Figure 4 Accuracy changes as the number of scales increases. Accuracy increases till a critical scale number after which it starts decreasing

3.2 Window Type

The window type for LPA filter can be chosen as uniform or Gaussian. Number of scales was set to 4 and both the window types were tested. The average accuracy with uniform window was 94.60 while with a Gaussian the average accuracy was 93.27.

3.3 Using Estimate and Derivative Together

A combination of an LPA estimate kernels along with the higher order LPA kernels which are differential filters can be used together to get the final feature vector. The higher order kernel used is the first derivative of the estimate. The average accuracy of the combination of these kernels was 92.24. Thus, a classification is more efficient when only the first derivative of the kernel estimates is used.

3.4 Block size

Block size is an important parameter as the performance of the classifier critically depends on it. The size of the block depends on the size of face image which is being processed. If block size is too small it will not be able to capture the relation between the surrounding pixels and if it is too big then it will no longer be able to capture the local information from the image. Thus, we varied the block size and observed how the accuracy changes with it. This experiment was done for images of different sizes. The size of images was varied from 35x35 to 55x55 with a step size of 5 pixels. It was observed across the images of different sizes that the best accuracy was achieved for the block size is approximately $1/8^{th}$ of the size of the face image. Thus in our further experiments we have fixed the image size to 40x40 and the block size is chosen as 5x5.

3.5 Yale Dataset

The performance of LPA-LBP-HS based methods was compared with Eigenface, Fisherface, LGBPHS and LBP. For Yale and ORL database LPA-LBP-HS based methods outperforms these methods. The results for Yale dataset are shown below. The images were resized to 40x40 size for all the tests and SVM linear function classifier was used for all. The average accuracy for the Yale database for these methods is shown in Figure 5. For the results shown in Figure 5, 8 images from each class were used in the training set and the rest of the images from each class were used as testing images.

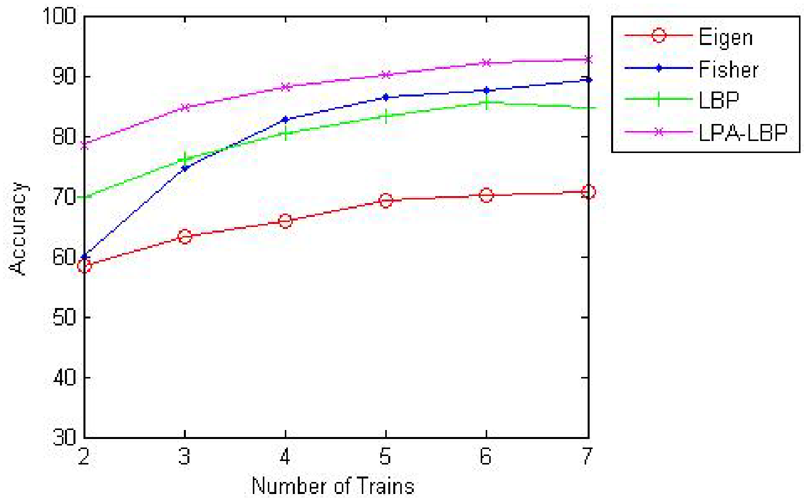

Figure 5. The figure shows how the performance of the different methods varies as the number of training samples are varied

Next the experiments were done by reducing the number of training samples for each class. The training samples were reduced from 7 to 2 and the results are shown in Figure 5. Better performance of LPA-LBP based method becomes prominent especially when the number of training samples is less. Conclusions that can be drawn from these observations are as follows

1. The proposed method when compared with the Gabor filters based face classification method gives much better result. In our method we are extracting directional information using LPA directional filters while Gabor method captures spatial localization, orientation characteristic based information. Thus, features based on directional information are more discriminative for the purpose of face classification.
2. The proposed method gives better results even when the number of training samples are considerably small. So even with a few training images critical features for a face are captured. Methods such as Fisherface perform well when the number of training images are high, but their performance reduces considerably with low number of training samples.

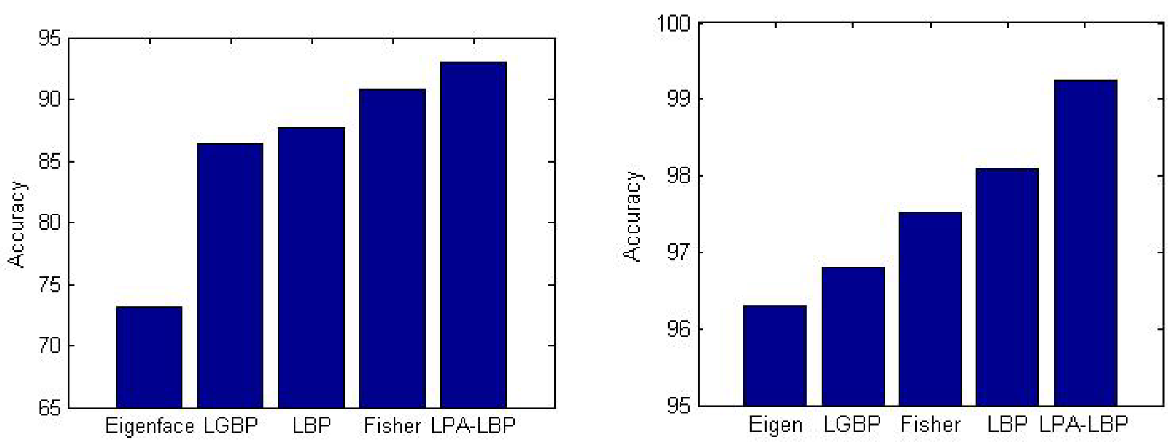

Figure 6. Accuracy for different methods is compared. In the figure on left 8 training samples were used on Yale dataset for each class while rest were used as testing samples. Figure on right compares the accuracies of different methods on ORL dataset by using 8 samples for training and rest as testing data.

3. Comparing the dimensionality and computational complexity of Gabor and LPA based method we can observe that Gabor based method extracts features from 5 scales and 8 orientations while in LPA we use 4 directions

and 4 scales. Thus, in Gabor based method 40 images are obtained while in our case number of directional faces is 16. Extracting features from 40 images can pose computational problems and also problem related to dimensionality. Down-sampling is generally applied in Gabor based methods to reduce dimensionality but it also results in a loss of some important features.

3.6 ORL dataset

Similar tests were done on ORL dataset. The dataset consist of 10 images of 40 individuals. For our test, the number of training images were set as 8 for each class and the rest of images were used for testing. The result of the proposed method is compared with those of Eigenface, Fisherface, LBP[4] and LGBP[5] methods and demonstrated in Figure 6. Thus, it can be seen that the face representation method based on LPA-LBP achieves better classification accuracy than other methods for both Yale and ORL dataset of images. Note, that even when the number of training samples are reduced, proposed method achieves betters results than existing methods.

4. DISCUSSION

This paper proposes a novel method for face representation through which we achieve better accuracy in face classification than other state-of-the-art methods. We have utilized LPA filter to extract directional information at multiple global scales. LBP and blockwise processing is done to capture the local and pixel level relations. Capturing information at different levels from a face image, we get a robust feature vector which is invariant to illumination and pose changes. Experiments done on Yale and ORL dataset of face images show that the proposed method outperforms existing methods such as Eigenface, Fisherface and LBP.

REFERENCES

[1] Turk, M. and Pentland, A. "Face recognition using Eigenfaces," in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, 586-591(1991)

[2] Belhumeur, P., Hespanha, J. and Kriegman, D. "Eigenface vs. Fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Anal and Mach Intell., 19, 711-720(1997).

[3] Martinez, A.M. and Kak, A.C. "PCA versus LDA," IEEE Transactions on Pattern Anal and Mach Intell, 23, 228-233(2001).

[4] Ahonen, T., Hadid, A. and Pietikainen, M., "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., 28, 2037-2041(2006).

[5] Zhang, W. C., Shan, S. G., Gao,W. and Zhang,H. M. "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A novel non-statistical model for face representation and recognition," Proc. IEEE Int. Conf. Computer Vision, 786-791(2005).

[6] Lei, Z., Liao, S., He, R., Pietikäinen,M. and Li, S. Z. "Gabor volume based local binary pattern for face representation and recognition," Proc. IEEE Int. Conf. Automatic Face & Gesture Recognition, 1-6(2008)

[7] Li, X., Hu, W., Zhang, Z., and Wang, H. "Heat Kernel Based Local Binary Pattern for Face Representation" Signal Processing Letters, IEEE, 17, 308-311(2010)

[8] Katkovnik, V., Foi, A., Egiazarian, K., and Astola, J., "Directional Varying Scale Approximation for Anisotropic Signal Processing," EUSIPCO Proc. of XII European Signal Process. Conf., 101-104(2004).

[9] Katkovnik, V., Egiazarian, K. and Astola, J., "Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule," J. of Math. Imaging and Vision 16, 223-235(2002).

[10] Katkovnik, V., Paliy, D., Egiazarian, K. and Astola, J., "Frequency domain blind deconvolution in multiframe imaging using anisotropic spatially-adaptive denoising," EUSIPCO Proc. 14th European Signal Process. Conf., (2006).

[11] Katkovnik, V., Astola,J., Egiazarian, K., "Phase local approximation (PhaseLa) technique for phase unwrap from noisy data," IEEE Trans. Image Process., 17, 833-846(2008).

[12] Paliy, D., Katkovnik, V., Bilcu, R., Alenius, S., Egiazarian, K., "Spatially Adaptive Color Filter Array Interpolation for Noiseless and Noisy Data," International Journal of Imaging Systems and Technology (IJISP), 17, 105-122(2007).