ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 10, No. 1, 2015, *pp*. 029-039

Fusing Shearlets and LBP Feature Sets for Face Recognition

Zhiyong Zeng [1]

[1] Faculty of Software, Fujian Normal University, Fuzhou, 350108, China

(*Received October 07, 2014, accepted December 24, 2014*)

Abstract. To aim at the challenge of face recognition to uncontrolled situations, robust face recognition system requires to take into account different kinds of face appearance feature. However, most existing methods only use features of just one type. We show that integrating two of global and local representations, Shearlets features and local binary pattern (LBP), which gets better performance than either alone. Shearlets features not only utilize scale and position information of different scales of decomposed image, but also use directional information. Shearlets features primary capture facial global attributes while LBP encode small local appearance details. Both feature sets are high dimensional so it is beneficial to apply block-based fisher linear discriminant (BFLD) and PCA to reduce dimensionality prior to normalization and integration. Then low dimensional Shearlets and LBP sets are combined by score level fusion. the proposed method is evaluated on two challenge face databases including MPIE and FERET with promising results.

Keywords: face recognition, Shearlets features, local binary pattern, BFLD, PCA, feature integration

1. Introduction

Face recognition (FR) is an active research issue in the area of computer vision and pattern recognition, FR has a wide range of applications, including information security, smart card, law enforcement, video surveillance and access control. However, how to extract effective feature representation to describe a face is critical for face recognition. In the past several decades, many face algorithms were proposed by researchers. Most of the appearance-based face recognition methods perform some kind of subspace analysis in the image space to extract the relevant feature vectors. The most widely used subspace analysis methods are Principal Component Analysis (PCA) [1], Linear Discriminant Analysis (LDA) [2], and a blind source separation technique, called Independent Component Analysis (ICA) [3]. However, the above methods are only designed for discovering the global features of data, while neglecting the local structure of the data. In fact, local features show certain robustness to local deformations of face images, for example, pose, expression and occlusion. Ojala presented one widely influential face descriptor, LBP descriptor, to facilitate rotation invariant analysis of image texture at multiple scales [4]. The operator uses arbitrary circular neighbor sets instead of eight neighbors, the number of samples as well as sample radius can vary. In addition, operators with different parameters can be combined to produce multiple scale descriptor of texture. LBP descriptor is used to encode small appearance details of face and greatly improve the performance of FR [5]. However, the sparse sampling exploited by LBP operator with large neighbor radius may result in inadequate representation of the face image and more sensitive to the noise. Zhang utilizes the robustness of Gabor feature to illumination and expression variations and proposed a histogram sequence using local binary pattern descriptor on Gabor magnitude map [6]. Above methods only apply either global features or local features, the most recent researches find that both global features and local features play an important rule for face recognition, researchers are studying the fusion methods of the two kinds of features. Many feature combination methods have been proposed [7-10]. The common used strategy of the fusion is weighting the feature extracted from different blocks [7-8]. In [7] and [8] the kernel linear discriminant was used to reduce the dimension of the histogram feature. Su uses discrete Fourier transform and Gabor wavelet transform to extract global and local features, then global and local features were integrated in both serial and parallel manner, the system accuracy is greatly increased [9]. Xie utilizes local Gabor XOR pattern to encode the Gabor phase, and apply block-based Fisher’s linear discriminant to reduce the dimensionality of the proposed descriptor and fuse the local patterns of Gabor magnitude and phase for face recognition [10]. Yan regard Gabor magnitude pictures as smooth surface and get a face representation method called Gabor surface feature by completely describing the shape of Gabor magnitude pictures [11].

Despite the great success of Gabor feature-based local feature FR methods, the Gabor transformations need high computational cost and storage space. The Gabor transformations of an face image need to be implemented at multiple scales and orientations. Therefore, the many convolutions and Gabor feature maps produce the high time and space complexity during the Gabor feature generation, which prevents its widely ultilization in practical applications.

Afterwards, Yang proposed a namely monogenic binary coding local feature extraction scheme, which decomposes an original face into three complementary components: amplitude, orientation and phase, then the monogenic variation in each local region and monogenic feature in each pixel were encoded, the method has significantly lower time and space complexity than the Gabor transformation based local feature methods [12].

Although multiresolution techniques like wavelets have been found very useful in analyzing the image contents, it is well known that wavelets have limited ability in expressing directional information. In order to overcome these limitations, a great number of multiscale geometric analysis methods such as curvelets [13], contourlets [14] and ridgelets [15], owned good characteristic such as locality, multiresolution, directionality and anisotropy, were proposed over the years, shearlet transform is a newly addition. Multiscale algorithms based on shearlets not only have good localization and compactly support in frequency domain, but also have directionality and anisotropy. With those properties, shearlets can effectively represent image geometrical information of edges, feature points and texture, it has been ultilized in image separation [16], image denoising [17], and image edge detection [18]. However, only a few work has been done to solve face recognition and pattern classification problems. For example, Qu proposed a facial expression recognition algorithm based on shearlet transform, which is a new image time-frequency analysis method and provides directionality and anisotropy [19]. However, only low frequency components in shearlet transform are extracted as face image feature, all high frequency components are wholely neglected. In fact, high frequency components include many useful information for face recognition. In addition, Danti proposed a facial texture feature representation based on shearlet transform [20-21], however, this method uses mean and covariance of shearlet transform coefficients as feature representation of a face image, which leads to cost computational complex and low face recognition rate. Borgi proposed a sparse coding augmented approach based on shearlet network, and designed a fusion step by PCA-based method using a refined model of belief function based on the Dempster-Shafer rule in the context of confusion matrices, this method is robust to the problem of a single training sample per subject [22].

This paper presents a new method for face recognition which combines shearlets features and LBP features. shearlet features describe the shape and appearance information over different scales, and LBP captures the small structural details of the face image. The remaining part of the paper is organized as follows: In section 2, the shearlets transformation is introduced. Section 3 presents the primary component analysis for shearlets features as global features, and BFLD for LBP features as local features for face recognition. Section 4 presents the strategy of integration of shearlet feature and LBP feature. And experiments and results analysis are conducted in section 5, followed by a small discussion, conclusion and future work in section 6.

2. Shearlet Transform

In the past decade years, researchers spent much effort to design various representation systems which sparsely approximate functions governed by anisotropic features such as prominent interest points, edges and borders in images. These representative systems include wavelets, curvelets, and contourlets. However, the above systems fail to provide a unified treatment of the continuum and digital world.

Wavelet representations are optimal for approximating data with pointwise singularities. However, wavelets are not very effective when dealing with multivariate data, wavelets cannot handle equally well distributed singularities such as singularities along curves because wavelets are isotropic objects and generated by isotropically dilating a single or finite set of generators. If distributed discontinuities such as edges of surface boundaries are present in two or higher dimension then wavelets fails to deal with such multivariate data. Curvelets can provide optimally sparse approximations of anisotropic features, but it has two drawbacks. Firstly, the curvelet is not singly generated, it is not derived from the action of countably many operators applied to a single or finite set of generating functions; secondly, its construction involves rotations and these operators do not preserve the digital lattice which prevents a direct transition from the continuum to the digital setting. Later contourlets were proposed by some researchers, but a proper continuum theory is missing in this approach.

Shearlets are derived from a single or finite set of generators and ensures a unified treatment of the continuum and digital world because the shear matrix preserves the integer lattice. It has some unique properties: it has a single or finite set of generating functions, it provides optimally sparse representation for a large class of multidimensional data, it is possible to use compactly supported analyzing functions, it has fast algorithmic implementations and it allows a unified treatment of the continuum and digital realms. Shearlets allows optimal encoding of several classes of multivariate data through its ability to sparsely represent anisotropic features and hence Shearlets emerged as the new tool for processing massive and higher dimensional data.

The Shearlets are an affine system which is parameterized by three parameters like- scaling, shear, and translation. The shear parameter can captures the direction of singularities of the face images. The direction of singularities to resolve the wavefront set of distributions can be precisely detected by continuous shearlets transform. The optimally sparse representations for 2-D functions that smooth away from discontinuities along curves can be provided by the discrete Shearlet transform. Another benefit of this approach is that, it provides a multi-resolution analysis similar to the one associated with classical wavelets, which is very useful for the development of fast algorithmic implementations.

The main idea for the construction of the Shearlet transform with discrete parameters for functions in $L^2(R^2)$ is the choice of a two-parameter dilation group, where one parameter ensures the multi-scale property and the second parameter provides a means to detect directions. Shearlets parameterize directions by slope rather than angles, so the shear matrix does preserve the structure of the integer grid, which is the key in enabling an exact digitization of the continuum domain Shearlets. Shearlet systems are designed to efficiently encode anisotropic features such as singularities concentrated on lower dimensional embedded manifolds. To achieve optimal sparsity, shearlets are scaled according to a parabolic scaling law encoded in the parabolic scaling matrix **A**, and exhibit directionality by parameterizing slope encoded in the shear matrix **B**, scale matrix A and shear matrix B is defined as:

$$A = \begin{bmatrix} a & 0 \\ 0 & \sqrt{a} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \tag{1}$$

Hence, Shearlet systems are based on three parameters: a > 0 is the scale parameter measuring the resolution level, and $s \in R$ is the shear parameter measuring the directionality, and $t \in R^2$ is the translation parameter measuring the position. This parameter space $R^* \times R \times R^2$ can be endowed with the group operation, which is defined as:

$$(a,s,t)(a',s',t') = \left(aa', s + s'\sqrt{|a|}, t + BAt'\right) \tag{2}$$

getting the so-called shearlet group B, which can be regarded as a special case of the general affine group. The Continuous Shearlet function arised from the unitary group representation is defined as:

$$\psi_{a,s,t}(x) = a^{-3/4}\psi\left(A^{-1}B^{-1}(x-t)\right) \tag{3}$$

For appropriate choices of the Shearlet $\psi \in L^2(R^2)$, the Continuous Shearlet Transform which is given as:

$$SH_f(a,s,t) = \langle f, \psi_{a,s,t} \rangle \tag{4}$$

is a linear isometry from $L^2(R^2)$ to $L^2(B)$. Alternatively, rather than defining the shearing parameter s on R, the domain can be restricted to, $|s| \leq 1$. This gives rise to the Cone-adapted Continuous Shearlet Transform, which allows an equal treatment of all directions in contrast to a slightly biased treatment by the continuous shearlet transform. By sampling the continuous shearlet transform on an appropriate discrete set of the scaling, shear, and translation parameters, it is possible to obtain a frame or even a parseval frame for $L^2(R)$.To obtain the discrete shearlets, three parameters are sampled as following:

$$a_j = 2^j (j \in Z), s_{j,k} = k a_j^{1/2} = k 2^{j/2} (k \in Z), and \qquad t_{j,k,m} = D_{a_j, s_j . k}(m \in Z^2) \tag{5}$$

The mother shearlet function ψ is chosen in a similar fashion as in the continuous case. The tiling of the frequency plane is shown in the figure 1. This system forms a parseval frame for $L^2(R)$, and they are optimally sparse. The discrete Shearlets on the cone, whose tiling of the frequency plane is shown in this figure 2 has the advantage that all directions are treated equally, also each scale is associated with a finite number of shear parameters. Shearlets can be easily constructed from separable and nonseparable generating functions, which is illustrated in figure 3. Shearlets generated from nonseparable functions cover the frequency plane more effectively and provide better frame bounds.

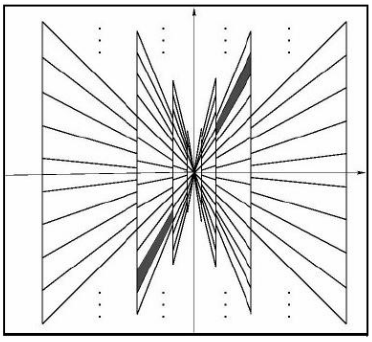

Fig. 1 The tiling of frequency plane

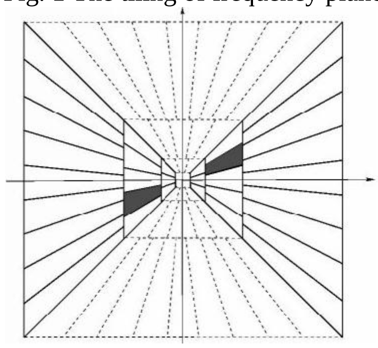

Fig. 2 The tiling of the frequency plane induced by discrete Shearlets.

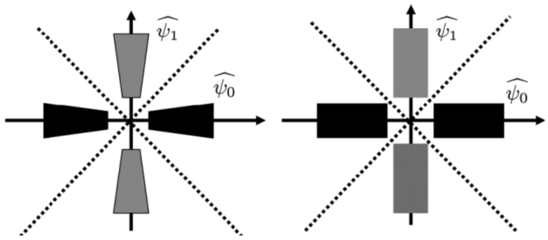

Fig. 3 Nonseparable Shearlets (left) and Separable Shearlets（right）

3. Global and Local Feature Extraction

The visual information of face can be captured by edges detected not only at different orientations but also multiple scales after a multi-scale decomposition of a face image has been performed by applying the shearlets transform. Generally, a face system includes two stages: training stage and classification stage. In training stage, each face image is decomposed along with horizontal and vertical directions by shearlet transform. Our decomposition is performed upto three levels by varying shearing parameters from $2^{(-j/2)}$ to $2^{(j/2)}$ and scale parameters from zero to two. So we obtain three horizontal cone coefficients and vertical cone coefficients. Then we perform PCA to the magnitude of these coefficients respectively, and then concatenate them into Shearlets features. we consider Shearlets features as the global features of the face image. Same procedure is repeated for all faces of the training database and a feature matrix is constructed from these low dimension Shearlets features. This feature matrix is used as the global features in classification stage to classify the unknown test face image. Shearlet features provide edge responses at a given scale and orientation because shearlet coefficients of large magnitude come from edges, so our Shearlets features are based on the magnitude of shearlet coefficients at different scales and orientations.

For clarity, we describe the PCA matrices learning procedure and feature extraction procedure in Algorithms 1 and 2, respectively.

Algorithm 1. Procedure of PCA matrices leaning

Input: *T*, the training set with *N* normalized face images.

Output: W, PCA matrices.

Step 1. for each image $I \in T$, shearlet decomposition is performed on *I*, and we get the shearlet coefficients, then we compute the magnitude of these coefficients $f_{s,d}, s = 1,2,3; d = 1,2$, where s is the scale index and d is the number of directionality on each scale.

Step 2. Based on $f_{s,d}$, these magnitude vectors are projected to a PCA subspace and then these transformed vectors are used to learn PCA matrices $w_{s,d}, s = 1,2,3; d = 1,2$.

Algorithm 2. Feature extraction using PCA matrices

Input: *I*, a normalized face image, the learned PCA matrices $w_{s,d}, s = 1,2,3; d = 1,2$.

Output: low-dimensional feature vector *F*.

Step 1. For image I, calculate its magnitude of shearlet coefficients $f_{s,d}, s = 1,2,3; d = 1,2$.

Step 2. Calculate their low-dimensional feature vectors $F_{s,d}, s = 1,2,3; d = 1,2$ respectively using linear transforms: $F_{s,d} = (W_{s,d})^T f_{s,d}$.

Step 3. Concatenate these low-dimensional feature vectors $F_{s,d}$ to form a face image feature representation *F*.

LBP descriptor is used to encode small appearance details of face. A typical LBP coding usually consists of the following thress stages: binary quantification, binary sequence generation, and binary sequence to decimal value conversion. In the procedure of LBP coding, the binary sequence generation and binary sequence to decimal value conversion are the common stages in most binary pattern coding algorithms. However, binary quantization usually depends on the property of feature map, and should be designed based on the physical meaning of the feature. Generally, LBP operator gives a label to every pixel of an image by thresholding the neighborhood of each center pixel value and using the results as a binary number. For example, given a pixel in a facial image, a LBP [4] code is calculated by comparing it with its neighbours:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & else \end{cases} \tag{6}$$

where g_c is the gray level of the central pixel, g_p is the gray value of its neighbors, *P* is the total number of involved neighbors and *R* is the radius of the neighborhood. Suppose the coordinates of g_c is (0, 0), then the coordinates of g_p are $(R*\cos(2*p/P), R*\sin(2*p/P))$. The gray values of neighbours that are not in the image grids can be estimated by interpolation. Suppose the image is of size $M \times N$. After the LBP pattern of each pixel is identified, a histogram is extracted to represent the facial image texture:

$$H(k) = \sum_{i=1}^{M} \sum_{j=1}^{N} f(LBP_{P,R}(i,j),k), k \in [0,L], f(x,y) = \begin{cases} 1, & x = y \\ 0 & else \end{cases} \tag{7}$$

where *L* is the maximal LBP pattern value. The *U* value of a LBP pattern is defined as the number of spatial transitions (bitwise 0/1 changes) in that pattern:

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \tag{8}$$

The uniform LBP patterns refer to the patterns which have limited transition or discontinuities $(U \leq 2)$ in the circular binary presentation [4]. In practice, the mapping from $LBP_{P,R}$ to $LBP_{P,R}^{riu2}$ (superscript "*u2*" means uniform patterns with $(U \leq 2)$, which has $P*(P-1)+3$ distinct output values, is implemented with a lookup table of 2*P* elements.

To achieve rotation invariance, a locally rotation invariant pattern could be defined as:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), if & U(LBP_{P,R} \leq 2) \\ P+1 & else \end{cases} \tag{9}$$

The mapping from $LBP_{P,R}$ to $LBP_{P,R}^{riu2}$ (superscript "*riu2*" means rotation invariant "uniform" patterns with $(U \leq 2)$, which has *P*+2 distinct output values, can be implemented with a lookup table.

Theoretically, LBP descriptor can be directly applied to face recognition by some similarity measurement and a classifier. However, it is not suitable because LBP feature dimension is very high due to the use of local block feature extraction. For example, a 128×128 facial image with 4×4 cells produces $32 \times 32 \times 59 = 60416$ -bins LBP feature vector. The dimensionality of this feature is very high, which will lead to cost computational complex and great storage space. In order to improve the efficiency and effectiveness of face recognition system, it is necessary that the technique of the dimensionality reduction is used. Therefore, we adopt the BFLD [10] scheme to reduce the LBP histogram feature dimension meanwhile enhancing its discrimination. LBP feature map is first divided into $M_b \times M_b$ blocks, and then each block is further divided into $M_r \times M_r$ sub-region. The histogram of the feature map in the sub-region is computed. Then all the histogram of the sub-region of each block are concatenated into the histogram feature of each block. We use LDA to project the histogram feature of each block on the subspace to learn the projection matrix from training set. Finally, we use this projection matrix to obtain the dimensionality reduced histogram feature of each block.

4. Integration of Shearlet and Local Feature Sets

Shearlets feature and LBP feature provide the complementary information to distinguish different human faces. To improve face recognition accuracy, we need to integrate these two kinds of features. Denote by F_i^p and F_i^g the dimensionality reduced LBP features in the i^{th} block of a probe face image and a gallery image, respectively, their similarity using cosine distance is defined as:

$$S\left(F_i^p, F_i^g\right) = \frac{\left\langle F_i^p, F_i^g \right\rangle}{\left\| F_i^p \right\| \times \left\| F_i^g \right\|} \tag{10}$$

where $\langle \bullet \rangle$ is the inner product operator. The similarity between the whole probe and gallery images is computed as:

$$S\left(F^p, .F^g\right) = \sum_{i=1}^{M_b \times M_b} S\left(F_i^p, F_i^g\right) \tag{11}$$

We can consider $S\left(F_i^p, F_i^g\right)$ as local classifier (LC), and then we compute the cosine distance of the dimensionality reduced Shearlets features and treat it as the global classifier (GC). Finally, we use the simple weighted sum to integrate the GC and LC into unified classifier (UC) as follows:

$$UC = w \times GC + (1 - w) \times LC \tag{12}$$

where w is the weight of the global classifier.

5. Experiment and Analysis

In this section, we first discuss parameter settings, and then we evaluate the proposed algorithm on a small scale benchmark datasets ORL and two large scale benchmark datasets: Multi-PIE [23] and FERET [24], which have been widely used to evaluate face recognition algorithms.

5.1. Parameter Settings

There are some parameters in different stages of the proposed algorithm, for example, multi-scale Shearlet transform, computing sub-region histogram and features integration. To consistent with other approachs and provide a fair comparison, no pre-processing is used in all the following experiments.

For all the datasets, to minimize the possible misleading results caused by the training data, the results have been averaged over five experiments, all conducted using the same parameters. Before the shearlets decomposition the images are resized to 128×128.

In this paper, the parameters in multi-scale shearlet transform are the number of scales, the number of orientations. The number of scales is often set to 3, and the number of orientations is often set to 6. when sub-regions LBP histogram are computed, the whole image will be partitioned into $M_b \times M_b$ blocks, then each block is further divided into $M_r \times M_r$ sub-regions. In each sub-region, the LBP histogram is extracted. In each block, the discriminative feature is extracted by BFLD. In the stage of fusing, two similarities LC and GC will be averaged by using the weight *w*. If no specific instruction, we fix $M_b = 4$, $M_r = 2$, $w = 0.45$, and the dimensionality of discriminative feature in each block as 200.

5.2. Experimental Results and analysis on ORL

We evaluated the proposed algorithm on ORL face dataset. ORL is a small face image datasets, which consists of 400 different images related to 40 individuals; each individual has ten different expression, viewport and illumination images. The ORL database contains only grayscale face images, these images are of size 92×112. In the procedure of PCA metrices learning, the training set is yielded by each class face images randomly selected five images, so the number of training samples is set *N*=200, the test set is produced by other images.

Fig. 4 Samples of ORL face datasets

In the experiment on ORL, LBP is used as baseline method. we compare the proposed method with the state-of-the-art algorithms, such as LGBP [6], HGPP [25], LGXP [10], MBC-F [12]. The face recognition rates are listed in Table I.

In our test on the ORL dataset, the proposed method outperforms LBP, LGBP and HGPP methods but is a little bit lower than LGXP and MBC-F methods. It is reason that the first three algorithms only use one type of feature and the latter two methods and the proposed method in this paper use fusion feature. These results demonstrate that the integration of shearlet and local feature sets is of higher discriminant and better generization than some traditional methods, it can overcome the affect of the variation of the face expression and position, and it can effectively eliminate the error of lighting variation, so are the LGXP and MBC-F methods.

Table I Recognition rate (%) comparison on ORL among different algorithms and proposed method

Methods	**ORL**
LBP	91.0
LGBP	96.3
HGPP	96.8
MBC-F	99.0
LGXP	99.3
Proposed method	98.9

5.3. Experimental Results and Analysis on Multi-PIE

We use large scale Multi-PIE to verify the performance of the proposed algorithm. The training set in this experiment consist of all the 249 subjects in Session 1. In order to make the face recognition more challenging, the probe set consists of four subsets with both illumination and expression variations in Session 1, 2 and 3. The seven frontal images with extreme illuminations {0,1,7,13,14,16,18} and neutral expressions are used for examples in the training set. Four frontal images with illuminations {0,2,7,13} and different expressions are used for examples in the testing set.

(a)

(b)

(c)

(d)

(e)

Figure 5: A subject in Multi-PIE database. (a) Training samples with only illumination variations. (b) Testing samples with surprise expression and illuminations in Session 2. (c) Testing samples with squint expression and illuminations in Session 2. (d) and (e) Testing samples with smile expression and illumination variations in Session 1 and Session 3, respectively.

In the experiment on Multi-PIE, LBP is also used as baseline method. we compare the proposed method with the state-of-the-art algorithms, such as LGBP [6], HGPP [25], LGXP [10], MBC-F [12]. The face recognition rates are listed in Table II. It is clear to see that LBP has the worst recognition rate. The proposed method has very competitive recognition rates as LGBP, HGPP, LGXP and MBC-F.

Table II Recognition rate (%) comparison on Multi-PIE among different algorithms and proposed method

Methods	Multi-PIE			
	Smile-S1	**Smile-S3**	**Surprise-S2**	**Squint-S2**
LBP	97.0	68.6	69.0	76.9
LGBP	99.2	77.6	72.1	83.4
HGPP	99.2	80.2	75.3	83.4
MBC-F	99.1	85.2	72.3	83.1
LGXP	99.2	77.8	72.5	83.4
Proposed method	99.0	82.5	73.6	83.2

In this experiment, it can be seen that the proposed method outperforms LBP, LGBP and LGXP method in all cases except Squint-S2 but is lower than MBC-F methods. These results demonstrate that the fusion of different global and local feature sets is of higher discriminant and better generization than only using one kind of feature, it can overcome the affect of the variation of the face expression and position, and it can effectively eliminate the error of lighting variation. There are two reasons: on the one hand, shearlet feature utilize directional information of Shearlets transform, on the other hand, after reduction, the combined feature sets not only describe the global shape of a face image at a wider range of scales, but also characterize the small detail of a face image.

5.4. Experimental Results and Analysis on FERET

We tested the proposed method on the FERET face database of fa, fb, fc, dup1 and dup2 images, The fa images are used as gallery images and the fb, fc, dup1 and dup2 images are used as probes. There are 1196 fa images, 1195 fb images, 194 fc images, 722 dup1 images and 234 dup2 images. All subjects (with 1 exception) have exactly one gallery and one probe image. The fa and fb images for a single subject vary only in expression (neutral versus smiling), fc images were taken under different lighting conditions, dup1 images were taken later in time, and dup2 images were taken at least a year after the corresponding gallery images. We use 1002 randomly selected images from the gallery set for training data.

Fig. 6 Some sample images of FERET database

In the experiment on FERET, LBP is still used as baseline method. we compare the proposed method with the state-of-the-art algorithms, such as Tan's method, LGBP [6], HGPP [25], LGXP [10], MBC-F [12]. The experimental results are listed in Table III.

From the experimental results in Table III, we can see that the proposed method outperforms LBP, LGBP, HGPP and Tan's method but is slightly lower than LGXP and MBC-F methods. These results demonstrate again that the information of direction base on shearlet, Gabor and monogenic transform is very useful, meanwhile, the fusion of different feature sets is of higher discriminant and better generization than LBP, LGBP, HGPP and Tan's method.

Meanwhile, in Table III, we can see that several algorithms have different recognition rates in the FERET dataset, all the methods have good recognition rate in the fb probe dataset, these results illuminate that the recognition rates of these algorithms are not affected by the variety of the face expression when the time of the gallery images taken is the same as the time of the probe images taken. The recognition rates of these algorithms except LBP almost remain unchanged for the fc probe dataset, the recognition rate of LBP method clearly decrease for fc probe dataset, which account for that these methods except LBP are insensitive to the difference of the illumination. Although the proposed method use PCA to reduce the dimensionality of shearlet feature, it has good recognition performance for fc probe dataset, this illustrates shearlets transform can eliminate the affection of variation of illumination.

Table III Recognition rate (%) comparison on FERET among different algorithms and proposed method

Methods	**FERET**			
	fb	**fc**	**dup1**	**dup2**
LBP	97.0	79.0	66.0	64.0
LGBP	98.1	97.0	73.9	71.1
HGPP	97.5	99.4	79.6	77.8
MBC-F	99.7	99.5	93.6	91.5
LGXP	99.0	99.0	94.0	93.0
Tan's method	98.0	98.0	90.1	85.0
Proposed method	99.3	99.0	91.2	90.5

The recognition rates of all the algorithms clearly decrease for the dup1 and dup2 datasets, but the worse results were got for the dup2 dataset. It is visible the longer time of the face image taken, the more variation of face image is owing to individual natural aging, especially local difference of face, it should lead to more affection for face recognition. In the dup2 dataset, the recognition rate of LBP is 64.0 percent, the recognition rates of LGBP and HGPP is 71.1 and 77.8 respectively, the recognition rates of other methods are more than 85.0 percent, the recognition rate of the proposed method is 90.5 percent, the integration of global and local feature sets could improve in average than LBP, in average than LGBP, in average than HGPP, respectively. this illustrates that the combination of the global and local feature with directional information are robust for the individual aging.

6. Conclusions

In the past decade years, face representation based on local statistical feature have attracted much attention and achieved great success in face recognition area. However, most local feature methods only use one kind of local structural information, they do not use global attribute of face image. So we investigated a novel face recognition method using reduced shearlets feature by PCA as global feature and reduced LBP feature by BFLD as local feature, we found that the integration of global and local feature sets is more potential features for the design of efficient face recognition system. Meanwhile, the unique features of shearlets transformation is optimally sparse representation for a large class of multidimensional data and unified treatment of the continuum and digital world, which might be helpful for improved recognition rate. The experimental results obtained on three different datasets are encouraging: we show that using the combined features based on decision fusion can be got good performance. The performance can be further improved by applying other strategy of features fusion and better classifiers.

7. References

[1] M. Turk, A. Pentland. “Eigenfaces for recognition,” Journal of Cognitive Nejroscience, 1991, 3(1):71-86.

[2] 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, 1997, 711–720.

[3] K. Baek, B. Draper, R. Beveridge, K. She. “PCA vs. ICA: A comparison on the FERET data set.” In: Joint Conference on Information Sciences, Durham, NC, March 2002, 824–827.

[4] T. Ojala, M. Pietikinen, T. Menp. “Multiresolution gray scale and rotation invariant texture analysis with local binary patterns[J]." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.

[5] T. Ahonen, A. Hadid, M. Pietikainen. “Face description with local binary patterns: application to face recognition[J].” IEEE Transactions on Pattern Analysis and Machine Intelligence , 2006 , 28 (12) : 2027-2041.

[6] W. C. Zhang, S. G. Shan, H. M. Zhang, J. Chen, X. L. Chen, W. Gao. “Histogram sequence of local Gabor binary pattern for face description and identification.” Journal of Software, 2006,17(12): 2508-2517.

[7] X. Tan, B. Triggs. “Fusing Gabor and LBP feature sets for kernel-based face recognition,” *IEEE International Workshop on analysis and modeling of faces and gestures*, 2007.

[8] X. Tan, B. Triggs. “Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions,” in *Proc. International Workshop on Analysis and Modeling of Faces and Gestures*, 2007, 168-182.

[9] Y. SU, S. G. Shang, X. L. Chen, W. Gao. “Integration of Global and Local Feature for Face Recognition.” Journal of Software, 2010, 21(8):1849-1862.

[10] S. F. Xie, S. G. Shan, X. L. Chen, J. Chen. “Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition,” *IEEE Trans. Image Processing*, 2010, vol. 19, no. 5, pp. 1349-1361.

[11] K. Yan, Y. B. Chen, D. Zhang. “Gabor Surface Feature for face recognition,” Proceedings of 2011 ACPR,288-292.

[12] M. Yang, L. Zhang, S. Shiu, D. Zhang. “Monogenic Binary Coding: An Efficient Local Feature Extraction Approach to Face Recognition,” *IEEE Trans on Information Forensics and Security*, 2012, 7(6): 1738-1751.

[13] X. X. Wu, J. Y. Zhao. Couvelet Feature Extraction for Face Recognition and Facial Expression Recognition. 6th International Conference on Natural Computation, Yantai, Shandong, China, Aug., 2010, 3:1212-1216.

[14] S. M. Lajevardi, Z. M. Hussain. Contourlet Structural Similarity for Facial Expression Recognition. 35th IEEE International Conference on Acoustics, Speech, and Signal Processing. Dallas, Texas, U. S. A., 2010, 1118-1121.

[15] E. J. Candes, D. L. Donoho. Ridgelets : A key to higher-dimensional intermittency. Phil. Trans. Roy. Soc. London

A357, 1999, 2495-2509.

[16] G Kutyniok, W. Q. Lim. Image separation using wavelets and shearlets. Curves and Surfaces, Lectures Notes in Computer Science, 2012, 6920 : 416-430.

[17] G. R. Easley, D. Labate, F. Colonna. Shearlet-based total variation diffusion for denoising. IEEE Trans. Image Process. 2009, 18(2):260-268.

[18] S. Yi, D. Labate, G. R. Easley, H. Krim. A shearlet approach to edge analysis and detection. IEEE Trans. Image Process, 2009, 18(5):929-941.

[19] Y. Qu, X. M. Mu, L. Gao, Z. W. Liu. “Facial Expression Recognition Based on Shearlet Transform”. Advances in Intelligent and Soft Computing, 2012, 159: 559-565.

[20] A. Danti, J. ,,K. M. Poornima. “Face Recognition Using Shearlets”. 2012 7th IEEE International Conference on Industrial and Information Systems, Chennai, Aug., 2012, 1-6.

[21] A. Danti, K. M. Poormima. “Feature Extraction Using Shearlet Cone Representation for Face Recognition”. Emerging Research in Electronics, Computer Science and Technology, Lecture Notes in Electrical Engineering, 2014, 248:571-578.

[22] M. A. Borgi, D. Labate, M. E. Arbi, C. B. Amar. “Shearlet Network-based Sparse Coding Augmented by Facial Texture Features for Face Recognition”, 17th International Conference Image Analysis and Processing, Naples, Italy, 2013, 8157: 611-620.

[23] R. Gross, I. Matthews. J. Cohn, T. Kanade, and S. Baker, “Multi-PIE,” *Image and Vision Computing*, 2010, 28: 807–813.

[24] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, “The FERET Evaluation Methodology for Face Recognition Algorithms,” *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 22, pp. 1090-1104, 2000.

[25] B. Zhang, S. Shan, X. Chen, and W. Gao, “Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition,” *IEEE Trans. Image Processing*, 2006, 16(1): 57-68.