Fusion of Zernike Moments and SIFT Features for Improved Face Recognition

Chandan Singh
Professor
Department of Computer Science, Punjabi University
Patiala, India

Ekta Walia
Asst. Prof., Department of Computer Science, South Asian University, New Delhi, India

Neerja Mittal
Asst. Prof., Department of CSE&IT, RBIEBT, Kharar, Distt. Mohali, India

ABSTRACT
Combining the feature sets that are invariant to global as well as to local variations of face images would be an efficient approach to construct an optimal face recognition system. Thus, identification and combination of complementary feature sets has become an active topic of research in recent days. In this paper, a combination of two useful methods, i.e. Zernike Moments (ZMs) and Scale Invariant Feature Transform (SIFT) has been proposed for the recognition of face images wherein the global information of face images has been effectively extracted by the ZMs approach while SIFT descriptor is used to locate local distinct keypoints. Exhaustive experiments are performed on ORL and Yale face databases. It has been observed that the proposed fusion achieves 98.5% and 91.67% recognition rates on ORL and Yale databases respectively. The inherent characteristics of ZMs and SIFT are retained in the combined descriptor and therefore the proposed approach is highly robust against pose, illumination and expression variations.

Keywords
Zernike moments (ZMs), Scale Invariant Feature Transform(SIFT), Invariant features, Global features, Local features, Face recognition.

1. INTRODUCTION
Face recognition is a technique that identifies the image of a person (query image) by comparing it with the previously recorded images in the database. It can also be used to verify a claimed identity. It has become an active area of research in the fields of computer vision, neuroscience, and psychology. Face recognition systems are no longer limited to the identity verification and surveillance tasks alone. Several tools have started utilizing the face recognition technology to realize the human intentions, actions and behavior for designing future smart environments [1]. This research area is growing day-by-day with significant improvements being proposed by the researchers working in this field. Despite, the availability of a number of face recognition approaches, the performance achieved by them is not so good in real world applications and also against the coarse variations in pose, illumination, expression, occlusion and noise [2].

Due to ample ongoing research in this area, there exist a significant number of feature extraction approaches in literature to represent the face images. These existing face recognition approaches are broadly classified in two categories - global feature extraction approaches and the local feature extraction approaches [3].

Global face recognition methods are based on statistical approaches wherein the features are extracted from whole face image. Most frequently used global methods include subspace based methods, spatial frequency techniques and moment based methods. Amongst the available face recognition approaches, moment based invariants are particularly useful because of their characteristics of providing compact face representation and being invariant to image rotation [4-6]. Some work using magnitude features of the ZMs in recognizing the face images is already done [7-9]. ZMs approach is observed to extract the global information of images more effectively than any other approach in this category. MPEG-7 uses some of these moment based approaches as region based shape descriptors for image retrieval [10-11]. Estimate of head movement by using the phase coefficients of ZMs of the original image and that of the rotated image is used to obtain feature sets found to be tolerant to pose variation [12-13]. The magnitude features of ZMs obtained at some higher order of moments are observed to be invariant to expression variation [14].

Local face recognition approaches deal with the interior information included within specific parts of the face images like the features of image patches containing eyes, nose, mouth, cheeks, etc. Recently, lots of work has been done on local feature extraction methods, because they have been proved to be robust against variations in illumination, facial expressions, noise and the occlusion. The local face recognition approaches have been classified in two categories. First, the sparse descriptor which initially divides the face images into patches and then illustrates its invariant features. Amongst these descriptors, SIFT, introduced by Lowe [15] consists of useful characteristics of being invariant to scale and rotation. SIFT descriptor has been effectively applied in some of the face recognition applications [16]. SIFT and multi-scale local binary patterns (MLBP) based local features have been used by Li et al. to perform age invariant face recognition [17]. A two-stage image matching scheme and a strategy of keypoint search for the nearest subject for the recognition of face images is presented by Geng et. al [18]. The local features extracted by Gabor filters are invariant to scale and orientation and are able to detect the edges and lines in the face images [19]. Second, the dense descriptors assess an image, pixel by pixel, in order to capture the finer local characteristics of face images. Due to its simplicity in extracting the local features, Local binary pattern (LBP) is one of the most widely used approaches in this category [20]. Several variants of LBP are available in literature for the representation of face images (with compact feature sets) as well as to improve the classification performance of this approach [21-22].

Though, a number of global and local methods have been devised for representation of face images, still no single approach is found to be suitable in most of the situations. Presently, researchers are working on the techniques that combine the global and local features together because the information conveyed by these two feature sets, is different. Specifically, global features are related to the general characteristic of whole face whereas local features describe the finer details inside some parts of face images so it seems logical to combine both of these features sets since the information conveyed by them belong to different attributes of the face images. The combined subspace based approach, using both global and local features obtained by applying Local Discriminant analysis (LDA) based method is proposed by Kim et al. in [23]. Fang et al. have proposed fusion of global Principal component analysis (PCA) features and Haar wavelet based local features for face verification [24]. Local and global information extracted by using the Discrete cosine transform (DCT) coefficients along with the Fisher classifier developed for high dimensional multi-class problem has been proposed in [25].The hierarchical ensemble of global and local information is presented by Su et al. in [26] to build a robust face recognition system.

Following the same opinion, in this paper we have presented the fusion of two effective feature sets, i.e. ZMs and the SIFT descriptor. The proposed approach, named ZMs+SIFT, focuses on the extraction of better and effective information for the recognition of face images by combining the inherent characteristics of both of the approaches. Both of the ZMs and SIFT approaches have a number of useful characteristics as described above. Hence, the fusion of these two approaches is believed to be robust against diverse variations in face images. The exhaustive experiments performed on popular ORL and Yale databases against scale, pose, illumination and expression variations have proved that the said hypothesis is correct. The results obtained on these databases show an improvement of approximately 10-15% in the recognition rate of the proposed ZMs+SIFT approach as compared to that of the ZMs and SIFT approaches individually.

The rest of the paper is organized as follows. Section 2 presents a brief overview of the ZMs and the SIFT approach. The procedure involved in the proposed fusion of ZMs and SIFT is described in Section3. Experiments and results obtained from the proposed method are presented in Section 4. Conclusions and future directions are presented in Section 5.

2. FEATURES UTILIZED

In this paper, we have used radial moment based ZMs features to obtain global features wherein the magnitude coefficients of moments are used as invariant image features. In order to capture the local features, distinct keypoints have been extracted by using the SIFT descriptor. The following subsections provide a brief overview of the methodology involved in the extraction of features.

2.1 Zernike moments (ZMs)

ZMs are the features generated by transforming the input image on a complex set of Zernike functions. For the computation of ZMs, initially the basis function of the ZMs is defined by computing its radial polynomials. Thereafter, the input image is projected onto this evaluated basis function of the ZMs. The Zernike basis function $V_{nm}(x,y)$ of order *n* having repetition *m*, within the unit disc are given by

$$V_{nm}(x,y) = R_{nm}(x,y)e^{jm\theta} \tag{1}$$

where $j=\sqrt{-1}$, $\theta = \tan^{-1}\left(\frac{y}{x}\right)$ (2)

and

$$R_{nm}(x,y) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s (x^2+y^2)^{\frac{n-2s}{2}} (n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \tag{3}$$

subject to the conditions $n \geq 0, |m| \leq n,$ and $n-|m| = even$. The angle θ lies between 0 and 2π and is measured with respect to x-axis in counterclockwise direction. The ZMs of order *n* and repetition *m* of a function $f(x,y)$ are defined by

$$Z_{nm} = \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 f(x,y)V_{nm}^*(x,y)dxdy \tag{4}$$

where $V_{nm}^*(x,y)$ is the complex conjugate of $V_{nm}(x,y)$. To compute the ZMs of a digital image, Zernike basis functions are defined within a unit circle, in which case the pixels located outside the circle are not involved in calculation. The discrete form of the ZMs of an image of size $N \times N$ pixels, is described as

$$Z_{nm} = \frac{n+1}{\pi}\sum_{\substack{i=0 \\ x_i^2+y_k^2\leq 1}}^{N-1}\sum_{k=0}^{N-1} f(x_i,y_k)V_{nm}^*(x_i,y_k)\Delta x_i \Delta y_k \tag{5}$$

where

$$x_i = \frac{2i+1-N}{D}, y_k = \frac{2k+1-N}{D}, \ and \ \Delta x_i = \Delta y_k = \frac{2}{D} \tag{6}$$

with

$$D = \begin{cases} N, & for\ inscribed\ circle\ within\ the\ image \\ N\sqrt{2}, & for\ outer\ circle\ containing\ the\ complete\ image \end{cases} \tag{7}$$

In this work, the value of D has been taken as $D = N\sqrt{2}$. The computational framework defined by Eq.(5) through Eq.(7) takes into account the image normalization and hence possesses the property of scale invariance on the assumption that the centre of the unit disc lies on the centroid of the image [6]. The image can be reconstructed by applying the inverse transformation as:

$$\hat{f}(x_i,y_k) = \sum_{n=0}^{n_{\max}}\sum_{m=-n}^{n} Z_{nm}V_{nm}(x_i,y_k) \tag{8}$$

where $n_{\max}$ is the maximum order of moments.

2.2 Scale invariant feature transform (SIFT)

The scale invariant feature transform (SIFT) algorithm, developed by Lowe [15] generates image features which are invariant to image translation, scaling, rotation and partially invariant to illumination changes. These features help in reliable matching between different views of the same object. SIFT features are extracted in the four stages [16] as described below:

1. **Scale space extrema detection**: The first step computes the locations of potential interest points (candidate keys) in the image by detecting the maxima and minima of a set of Difference of Gaussian (DoG) filters applied at

different scales all over the image. Given a Gaussian-blurred image,

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \qquad (9)$$

where $I(x,y)$ is the given image and * designates convolution operation.

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}} \qquad (10)$$

In order to efficiently detect the stable keypoints in scale and space, this method makes use of the scale-space extrema in the difference-of-Gaussian function convolved with the image, $D(x,y)$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y)$$

or $$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \qquad (11)$$

2. **Keypoint localization**: The detailed model is fit at each candidate key to evaluate its location and scale and these locations are refined by discarding points of low contrast.
3. **Assignment of orientation**: Thereafter, an orientation is assigned to each key point based on local image features.
4. **Keypoint descriptor**: Finally, a local feature descriptor is computed at each key point. This descriptor is based on the local image gradient transformed according to the orientation of the key point to provide orientation invariance. Every feature is a vector of dimension 128 distinctively identifying the neighborhood around the key point.

3. PROPOSED APPROACH

The proposed approach of combining the global ZMs features with that of local SIFT descriptor comprise of a number of useful characteristics. Among the various global shape descriptors, ZMs are observed to be one of the best shape descriptors because of their many attractive characteristics [7]. The prominent characteristics include orthogonal kernel functions that ensure minimum information redundancy between moments, rotation invariance exhibited by magnitude of ZMs, scope of easily achieving translation and scale invariance and robustness to image noise. Further, the phase coefficients of ZMs of the original image and that of the rotated image may be utilized to estimate angle of rotation between them [12-13]. The magnitude features of ZMs obtained at some higher orders of moments are invariant to changes in the expression [14]. On the other hand, local features extracted with SIFT keypoint detection are highly distinctive wherein only three keypoints are able to discriminate the images effectively. The keypoints extracted by the SIFT descriptor are also highly robust against scale and rotation variations [15]. Hence, the fusion of these complementary feature sets is observed to be invariant against diverse types of variations present in the face images.

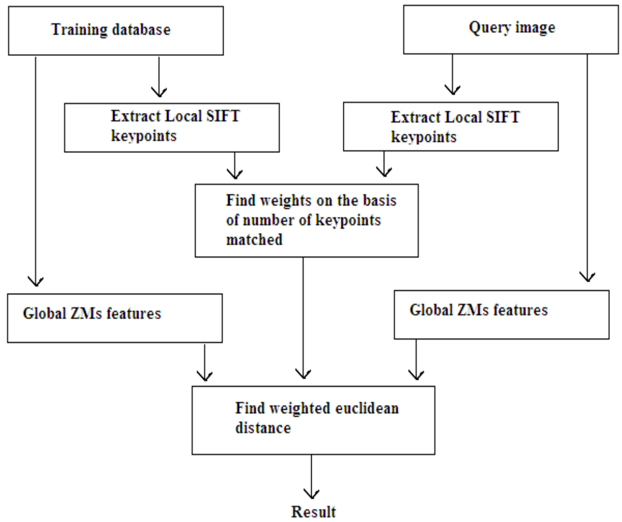

Fig. 1: Proposed Approach for fusion of ZMs and SIFT features

The outline of the procedure involved in combining the ZMs and SIFT approaches is depicted in Fig. 1. The face database consist of both the ZMs features and the SIFT keypoints whereas the said features for the query image are extracted online. For a given query face image, firstly the global ZMs features are extracted which are compared against the global features of the database images by using the nearest neighbor classifier (Euclidean distance). Thereafter, the numbers of keypoints of query image that match with that of the database images are found. The Euclidean distance is then weighted according to the number of keypoints that match.

4. EXPERIMENTS AND RESULTS

In order to evaluate the performance of the proposed combined ZMs+SIFT approach in comparison to that of the individual ZMs and SIFT approaches, the experiments are performed on two well known and calibrated face databases, namely, Yale face database [27] comprising of illumination and expression variations and the ORL face database [27] having small pose (tilt/yaw) changes. It is well known that the accuracy of the face recognition system is significantly affected by the kind of variations present in images of the face database as well as by the number of images of each person kept in the training set. Thus, exhaustive experiments are performed with respect to different types of variations present in the face images of these databases as well as by varying the number of training images of each person. Best results are highlighted in boldface. All the experiments are performed in Visual C++6.0 and MATLAB 7.0 under Microsoft Windows environment on a PC with 3.0 GHz CPU and 3 GB RAM.

4.1 ORL database

The ORL face database consists of a total of 400 images of size 112×92 pixels of 40 persons with 10 images per person in different states of variations. All the face images in this database are taken against a dark homogenous background. These images contain slight pose variation (tilt and yaw) upto ±20° with some basic facial expressions (smiling/not smiling, open/closed eyes). Sample images for one person from this database are shown in Fig. 2

Fig. 2: Sample face images for one person from ORL database

First of all, the experiments are performed by taking one image of each person in the training set and all of the remaining ones are used to formulate the test set. This setup is repeated for 10 times by taking every time a different image of each person in the training set. The recognition result over the said 10 different groups of training and test sets are presented in Table 1 depicted on the last page. From the results, it has been observed that the performance (on ORL) of SIFT descriptor is better than that of the ZMs against scale and pose variation. This demonstrates the good invariance characteristic of SIFT descriptor. Despite this, highest recognition results are achieved by the proposed ZMs+SIFT approach over all of the groups of training and test sets.

Secondly, the experiments are performed by taking different number of images in the training and the test set. For the said purpose, out of the ten images of each person in this database, first 2, 3, 4 and 5 images of each person are taken in the training set and the corresponding remaining 8, 7, 6 and 5 images are placed in the test set. The recognition results of each group of the training and the test sets for the considered approaches are presented in Fig. 4. It is clear from the presented results that superior recognition rate is achieved by the proposed ZMs+SIFT approach on all of the groups of the training and test sets. For example, in case when first three images of each person are taken in the training set while the remaining all are placed in the test set, the recognition rate of SIFT and ZMs approaches is 78.21% and 88.57%, respectively. On the other hand, superior recognition rate of 95.0% is achieved by the proposed ZMs+SIFT approach on the same set of training and test images.

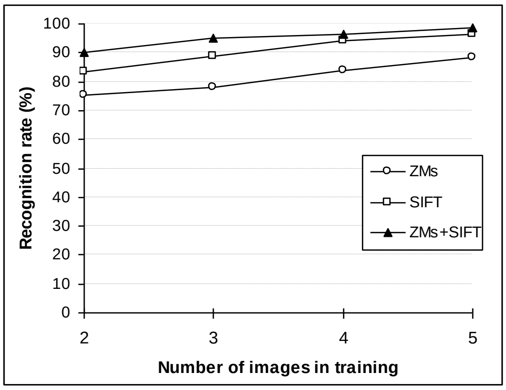

Fig. 4: Performance (%) on ORL database

4.2 Yale database

The Yale face database contains 11 images per person for 15 individuals resulting to a total of 165 images. The images in this database reveal major variations of illumination changes, different facial expressions and the persons wearing eyeglasses/no eyeglasses. The original size of the images in this database is 243×320 pixels with 256 gray levels. For the experiments, these are cropped down to formulate the face images of size 64×64 pixels. Sample cropped images from this database, for one person, are shown in Fig. 3.

Fig. 3: Sample cropped images for one person from Yale database

In first set of experiments, out of total 11 images of each person, one is taken in the training set and the remaining all are placed in the test set. The said process is repeated 11 times by taking different face image of each person in the training set. Table 2 represented on the last page, shows recognition results over 11 different runs of said setup of training and test sets. From the results presented, it has been observed that in most of the cases, the performance of the combined ZMs+SIFT approach is better than that of the performance of individual ZMs and SIFT approaches. In case the image containing coarse illumination (side lighting) is taken in the training set, the recognition rate of the proposed ZMs+SIFT approach is found comparable to that of the ZMs approach. SIFT keypoints are highly robust against scale and rotation variations. However, these are partially invariant to the illumination variation which results in the lower performance of the SIFT descriptor on this database.

Further, the experiments are performed by increasing the number of images per person in the training set. For this purpose, the first 2, 3, 4, 5 and 6 images of each person are taken in the training set while the corresponding remaining 9, 8, 7, 6 and 5 images are placed in the test set. The recognition results for each group of the training and the test sets, for the considered approaches are presented in Fig. 5.

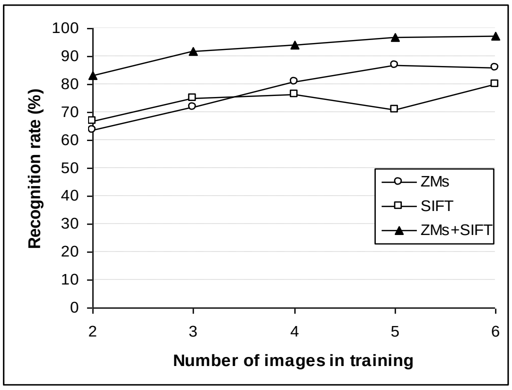

Fig. 5: Performance (%) on Yale database

From the results obtained, it is observed that approximately 10-15% improvement in the recognition results is achieved by the proposed fusion as compared to the individual approaches. In case of individual approaches, performance of SIFT descriptor is better when first 2 and 3 images of each person are taken in the training set whereas on other groups, i.e. on taking first 4, 5 and 6 images in the training set, ZMs approach performs better. In all cases, better recognition rate is achieved by the proposed ZMs+SIFT approach. Thus the proposed fusion of two complementary feature sets leads to improved method that is capable of capturing the gross characteristics as well as invariant details of the face images.

4.3 Noise variation

In order to analyze the performance of the proposed ZMs+SIFT approach in presence of additive noise, we add impulsive noise commonly named salt-and-pepper or spike noise to the face images of both the databases. In presence of spike noise, an image has dark pixels in bright regions and white pixels in dark regions [14]. A noise density of 0.05 is added in the images of test set and the training is done on original images, i.e. on images with no noise. Gaussian noise is one of the common noise types that occur due to variation in camera settings while capturing these images. Hence, the experiments are also performed by adding the Gaussian noise (zero mean and 0.01 variance) on images of these databases.

Table 3. Performance (%) against noise variation

Method	ORL database		Yale database	
	Spike	Gaussian	Spike	Gaussian
ZMs	85.0	85.5	82.22	83.33
SIFT	88.5	88.5	56.67	60.0
ZMs+SIFT	**96.0**	**96.5**	**94.44**	**93.5**

Herein this set of experiments, the first five images of each person from the ORL and the Yale databases are used in their respective training sets whereas the corresponding remaining images (with additive noise) are used for testing. Recognition rates for the considered approaches are presented in Table 3. From Table 3, it can be observed that the recognition rate of SIFT descriptor drops very sharply due to noise, which clearly implies that it is less invariant to image noise. On the other hand, the ZMs approach is significantly invariant to image noise. It can however be noticed that the proposed ZMs+SIFT method is also robust against the noise variation.

Table 4: Performance (%) comparison of some recent techniques on Yale and ORL databases

Methods	Yale database	ORL database
PCA [28]	53.3	90.5
Two-Dimensional PCA [28]	73.3	86.0
Feature fusion approach [28]	82.7	92.0
Combined feature Fisher classifier (CF^2C) [25]	-	95.1
Fisherface [29]	63.3	92.0
Intrinsicface [29]	68.3	95.0
ZMs+Optimal Similarity Measure (ZMs+OSM) [10]	79.67	89.85
Pseudo ZMs+OSM [10]	80.7	95.67
Complex Zernike moments [12]	-	90.5
Component features of ZMs (RICZMs)[13]	-	92.0
ZMs	71.67	88.5
SIFT	75.0	96.5
ZMs+SIFT	**91.67**	**98.5**

4.4 Performance comparison with some recent methods

Herein this subsection, we have compared the performance of the approaches considered in this study against some well established and recent approaches in the field of face recognition. Table 4 shows the recognition accuracy (in percentage) of some well known and recent methods on Yale database (for first three images of each person in training) and ORL databases (for first five images of each person in training). From this, it is observed that the recognition accuracy of the proposed ZMs+SIFT approach is superior as compared to other approaches on ORL and Yale databases.

5. CONCLUSION

In this paper, we have proposed the fusion of two effective feature sets, i.e. the global ZMs approach and local SIFT descriptor. These approaches, when applied individually are observed to provide good recognition performance on the face images containing some definite variations. The fusion of the said approaches includes the benefits of both of them and as such proves to be invariant against various types of variations present in face images. From the experimental results performed on ORL and Yale face databases, it is observed that the proposed ZMs+SIFT approach exhibit an improvement of 10-15% in the recognition rate when compared to the performance of these approaches individually. It has also been observed that the combined ZMs+SIFT approach generates superior results against the scale, pose, illumination and expression variations. It is further found to be robust against image noise.

6. REFERENCES

[1] Zhang, X. and Gao, Y. 2009 Face recognition across pose: A review. Pattern Recognition, 42: 2876-2896.

[2] Abate, A.F., Nappi, M., Riccio, D. and Sabatino, G. 2007 2D and 3D face recognition: A survey. Pattern Recognition, 28: 1885-1906.

[3] Hjelmas, E. and Low, B.K. 2001 Face detection: A survey. Computer vision and Image Understanding, 83: 236-274.

[4] Teague, M. R. 1980 Image analysis via the general theory of moments. Journal of Optical Society of America, 70: 920-930.

[5] Singh, C. 2006 Improved quality of reconstructed images using floating point arithmetic for moment calculation. Pattern Recognition, 39: 2047-2064.

[6] Wee, C–Y. and Paramesran, R. 2007 On the computational aspects of Zernike moments. Image and Visual Computing, 25: 967-980.

[7] Nor'aini, A. J., Raveendran, P. and Selvanathan, N. 2006 Human face recognition using Zernike moments and nearest neighbor classifier. In Procedings of 4th student Conf. on Research and Development, 27-28 June Selangor: 120-123.

[8] Nor'aini, A. J., Raveendran, P. and Selvanathan, N. 2006 A comparative analysis of Zernike moments and Principal Components Analysis as feature extractors for face recognition. In Proceedings of 3rd Kuala Lumpur International Conf. on Biomedical Engineering: 37-41.

[9] Singh, C., Walia, E. and Mittal, N. 2011 Magnitude and phase coefficients of Zernike and Pseudo Zernike moments for robust face recognition. In Proceedings of the IASTED International Conf. on Computer Vision (CV 2011),1-3 June Vancouver, BC, Canada: 180-187.

[10] Zhang, D. and Lu, G. 2004 Review of shape representation and description techniques. Pattern Recognition, 37: 1-19.

[11] Zhang, D. and Lu, G. 2003 Evaluation of MPEG-7 shape descriptors against other shape descriptors. Multimedia Systems, 9: 15-30.

[12] Singh, C., Mittal, N. and Walia, E. 2011 Face Recognition using Zernike and Complex Zernike Moment Features. Pattern Recognition and Image Analysis, 21(1): 71-81.

[13] Singh, C., Walia, E. and Mittal, N. 2011 Rotation Invariant Complex Zernike Moments Features and their Application to Human Face and Character Recognition. IET Computer Vision, 5 (5): 255-265.

[14] Lajevardi, S.M. and Hussain, Z.M. 2010 Higher order orthogonal moments for invariant facial expression recognition. Digital Signal Processing, 20: 1771-1779.

[15] Lowe, D. G. 2004 Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60 (2): 91-110.

[16] Bicego, M., Lagorio, A., Grosso, E. and Tistarelli, M, 2006 On the use of SIFT features for face authentication.

In Proceedings of the International IEEE workshop on Biometric Authentication, IEEE Computer Society Press.

[17] Li, Z., Park, U. and Jain, A.K. 2011 A Discriminative Model for Age Invariant Face Recognition. IEEE Transactions on Information Forensics and Security, 6 (3): 1028-1037.

[18] Geng, C. and Jiang, X., 2011 Face recognition based on the multi-scale local image structures. Pattern Recognition, 44: 2565-2575.

[19] Huang, L., Shimizu, A. and Kobatake, H. 2005 Robust face detection using Gabor filter features. Pattern Recognition Letters, 26 (11): 1641-1649.

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

[21] Ahonen, T., Hadid, A. and Pietikäinen, M. 2006 Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28 (12): 2037-2041.

[22] Tan, X. and Triggs, B. 2010 Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. IEEE Transactions on Image Processing, 19 (6): 1635-1648.

[23] Kim, C., Oh, J. and Choi, C. 2005 Combined Subspace Method Using Global and Local features For Face Recognition. In Proceedings of IEEE International Joint Conf. on Neural Networks, 31 July Montreal Canada, 4: 2030-2035.

[24] Fang, Y., Tan, T. and Wang, Y. 2002 Fusion of Global and Local Features for Face Verification. In Proceedings of IEEE Conf. on Pattern Recognition, 2: 382-385.

[25] Zhou, D., Yang, X., Peng, N. and Wang, Y. 2006 Improved-LDA based face recognition using both facial global and local information. Pattern Recognition Letters, 27: 536-543.

[26] Su, Y., Shan, S., Chen, X. and Gao, W. 2009 Hierarchical Ensemble of Global and Local Classifiers for Face Recognition. IEEE Transactions on Image Processing, 18 (8):1885-1895.

[27] Olivetti Research Laboratory (ORL) and Yale face databases http://www.face-rec.org/databases/

[28] Xu, Y., Zhang, D., Yang, J. and Yang, J. –Y. 2008 An approach for directly extracting features from matrix data and its application in face recognition. Neurocomputing, 71: 1857-1865.

[29] Wang, Y. and Wu, Y. 2010 Face recognition using Intrinsicfaces. Pattern Recognition, 43: 3580-3590.

Table 1: Recognition rate on taking one (different) image in training from ORL database

Method	Image Number/Type of variation (shown on sample images of one person)									
	1	2	3	4	5	6	7	8	9	10
	 neutral	 left	 scale	 right	 left	 right	 Up	 scale	 up/scale	 down/scale
ZMs	64.44	67.78	69.44	66.94	63.61	66.11	66.67	71.11	66.67	61.67
SIFT	74.72	72.5	71.39	73.06	73.06	71.39	69.72	73.89	71.39	72.5
ZMs+SIFT	81.94	81.39	82.78	82.78	80.22	78.6	79.72	84.17	83.33	88.33

Table 2: Recognition rate on taking one (different) image in training from Yale database

Method	Image Number/Type of variation (shown on sample images of one person)										
	1	2	3	4	5	6	7	8	9	10	11
	 neutral	 surprise	 wink	 wearing eyeglasses	 smiling	 side lighting	 front lighting	 front lighting	 side lighting	 expression	 closed eyes
ZMs	43.33	61.33	73.6	18.0	68.67	64.0	24.67	68.67	72.67	66.67	66.0
SIFT	32.67	58.67	67.3	67.33	64.0	37.3	64.67	72.0	33.67	72.0	75.33
ZMs+SIFT	51.33	78.0	84.0	70.67	80.0	64.0	67.33	89.33	72.0	86.67	90.0