Proceedings of the International Conference on Machine Vision and Machine Learning
Prague, Czech Republic, August14-15, 2014
Paper No. 125

Face Recognition in Low-resolution Images by Using Local Zernike Moments

Tolga Alasag, Muhittin Gokmen
Department of Computer Engineering, Istanbul Technical University
ITU Ayazaga Campus 34469 Maslak, Istanbul, Turkey
talasag@itu.edu.tr;gokmen@itu.edu.tr

Abstract **-** In this paper, we propose a method that uses Local Zernike Moments (LZM) for face recognition in low-resolution face images. Global Zernike Moments produce one moment value for whole image, whereas LZM are based on the evaluation of the moments for each pixel. LZM are shown to achieve significant success and robustness in face recognition. In order to further increase the robustness of LZM against low resolution and blurred face images, we use a scale space representation similar to scale-invariant feature transform (SIFT) algorithm. The performance of LZM in low-resolution face images is tested on FERET database. Then a face recognition framework is formed according to these tests and results show that the proposed framework is promising for real world applications.

Keywords**:** Face Recognition, Local Zernike moments, Low resolution images.

1. Introduction

In recent years, a great deal of methods for face recognition in general scenes have been proposed. Although the proposed methods can recognize faces at high precision and high speed (Jafri and Arabnia, 2009), faces in real images such as the ones from the security and surveillance systems are not always sufficiently clear. Therefore, the face recognition methods must succeed in low resolution, blurriness, variation of poses, illumination conditions and partial occlusions.

Most of the existing methods, such as (Zhou et al., 2011), try to solve the low-resolution face problem by generating a space representation between low and high resolution face images. (Biswas et al., 2011) consider pose and illumination variations using a space representation.

The other group of methods are based on super resolution algorithm. However these methods have performance problems in real world applications. Bo and Hong propose coupled locality preserving mappings method to get rid of super resolution calculations (Li et al., 2010). The other approach proposed by (Hennings-Yeomans et al., 2008) uses multiple camera images and super resolution models

In this study, we propose a method that uses LZM and Gauss scale space for face recognition in low-resolution images in which the robustness of LZM against low resolution and blurred face images is increased.

The performance of LZM in low-resolution face images is examined on FERET database. Several tests are carried with various settings and parameters. Then a face recognition framework is designed according to the test results. Block diagram of the test bed and the framework is shown in Fig. 1.

The remainder of the paper is organized as follows. Sections 2-5 briefly describe the methods used in the proposed algorithm, such as LZM and Gauss scale space. Experimental results on FERET database are given section 6. Finally section 7 concludes the paper.

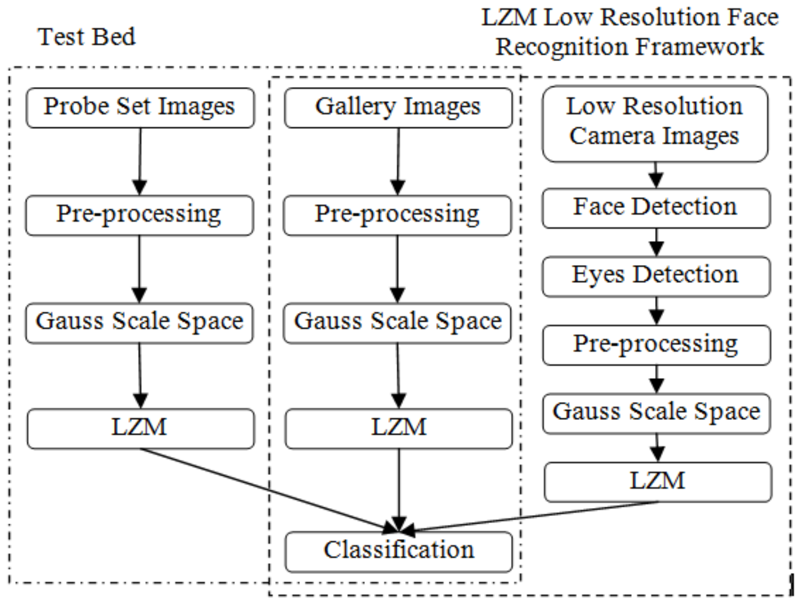

Fig. 1. Test bed and the proposed framework for face recognition from low-resolution images.

2. Local Zernike Moments

Zernike moments are based on the calculation of the complex moment coefficients and are successful in character recognition of images that contain distinctive shape information like characters (Khontanzad and Hong, 1990). However these holistic moments were seen inadequate for the face images and for this reason a novel face representation method called Local Zernike Moments (LZM) was proposed and is shown to be successful in face recognition (Sariyanidi et al., 2012). The LZM method localizes the calculation of the moments around for each pixel. As a result, a complex moment image is estimated for each moment component. Final feature vector is obtained by concatenating the extracted phase-magnitude histograms at each sub region that is formed by dividing each moment image to non-overlapping sub regions (Sariyanidi et al., 2012).

Local shape variations in the low resolution face images are very important for face recognition. For this reason, LZM are used as local phase-magnitude histograms (H-LZM).

2. 1. Transformation of Local Zernike Moment

The purpose of using LZM transformation in face images is to stimulate the local shape characteristics and to describe the local shape statistics of the transformed images (Sariyanidi et al., 2012). The LZM transforms are defined as

$$Z_{mn}^{k}(i,j) = \sum_{p,q=-\frac{k-1}{2}}^{\frac{k-1}{2}} f(i-p, j-p) V_{nm}^{k}(p,q) \tag{1}$$

$$V_{nm}^{k}(i,j) = V_{nm}(p_{ij}, \theta_{ij}) \tag{2}$$

where V_{nm}^{k} express the $k \times k$ sized filter kernels of Zernike polynomial and $V_{nm}(p_{ij}, \theta_{ij})$ express an orthogonal set of Zernike polynomials given by

$$V_{nm}(p,\theta) = R_{nm}(\rho)e^{jm\theta} \tag{3}$$

where

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s(\rho)^{n-2s}(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} . \tag{4}$$

Different number of images are calculated according to the order of the moment using (1)-(4). The number of active moments are defined by

$$K(n) = \begin{cases} \frac{n(n+2)}{4} & \text{if } n \text{ is even} \\ \frac{(n+1)^2}{4} & \text{if } n \text{ is odd} \end{cases} . \tag{5}$$

2. 2. Extraction of Feature Vector

After the application of LZM to face images, the next step is the production of the feature vector from the moment components. Dividing the face images into sub regions has an important role in getting better results in terms of recognition capability of the feature vector. Moment components are divided into non-overlapping sub regions, first into $N \times N$ sub regions and then into a half block shifted $(N-1)\times(N-1)$ sub regions that have the same size as the previous ones. Finally the phase-magnitude histograms are calculated using each of these sub regions.

Calculation of the phase-magnitude histograms is based on the angle and phase of each pixel. These histograms are divided into the angle interval of $[0,2\pi]$ that contain b bins. The magnitude at each pixel is added to the bin corresponding to the phase of the same pixel. After phase-magnitude histograms are evaluated for each sub region by this way, they are normalized to have unit norm. Then the histograms are concatenated to obtain the final feature vector of the input image.

In this work, all LZM tests are performed on the FERET face database (Rizvi et al., 1998). The database is formed with training images called gallery and four groups of test images called probe set. While gallery contains 1196 face images, probe sets FaFb, FaFc, Dup1 and Dup2 contain 1195, 194, 722 and 234 images, respectively.

Table 1 shows the recognition ratios of LZM with different settings on FERET database (Sariyanidi et al., 2012).

Table. 1. FERET database recognition rates of LZM (Sariyanidi et al., 2012)

Method	FaFb	FaFc	DupI	Dup II
H-LZM	95.0	87.1	73.8	70.9
H-LZM (Weighted)	97.5	95.4	78.5	76.1
H-LZM^2-I	96.2	97.9	79.6	76.9
H-LZM^2-I (Weighted)	98.7	99.5	83.9	82.5
H-LZM^2-R	96.2	96.9	77.4	73.5
H-LZM^2-R(Weighted)	98.7	99.0	83.2	81.2
H-LZM^2-IR	96.3	97.9	79.9	76.5
H-LZM^2-IR (Weighted)	98.7	99.5	84.8	82.5

3. Scale Space for Face Representation

In order to increase the robustness of LZM against low resolution and blurred face images further, a scale space algorithm like SIFT algorithm is used (Lowe, 2004).

The Scale space, being different than the Gauss pyramid, is formed by a particular number of images per level. Each level has a group of images that are formed step by step with convolution by increasing σ value of the Gaussian filter at each step and then the last image is down-sampled (Lowe, 2004).

To build a scale space, several parameters should be determined including the number of levels called octaves denoted by o, the number of images per octaves called octave layers denoted by s, the input Gauss sigma value σ and the input image I.

I is convolved s times for the first octave. So, first image of the octave $(n=0)$ is blurred by σ and last image of the octave $(n=s)$ is blurred by 2σ. Convolving an image by Gaussians for a number of times is equivalent to convolving it once by a single Gaussian with variance equal to the sum of all the variances of the other Gaussians as shown in (6), (7) and (8).

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 \tag{6}$$

For $n \geq 1$ and for each image;

$$std(n) = \sigma(n-1)\sqrt{k^2-1},\ k = 2^{1/s},\ \sigma(n) = \sigma k^n \tag{7}$$

or briefly

$$std(n) = \sigma k^{n-1}\sqrt{k^2-1}\,. \tag{8}$$

$s+1$ images are created per octave to generate a Gauss scale space as follows

$$\begin{aligned}
&blur[0] && = I * G(\sigma) \\
&blur[1] = blur[0] * G(std(1)) && = I * G(k\sigma) \\
&blur[2] = blur[1] * G(std(2)) && = I * G(k^2\sigma) \\
&\dots && \\
&blur[s] = blur[s-1] * G(std(s)) = I * G(k^s\sigma) = I * G(2\sigma) &&
\end{aligned} \tag{9}$$

where operator $*$ represents convolution.

To generate Gauss scale space for all octaves, primarily first octave is calculated from I and then its 2σ blurred image is down-sampled. Thus effective Gaussian sigma value increases: the first octave goes from σ to 2σ, the second octave from 2σ to 4σ, and so on. This process continues for all of the octaves. A sample Gauss scale space is shown Fig. 2.

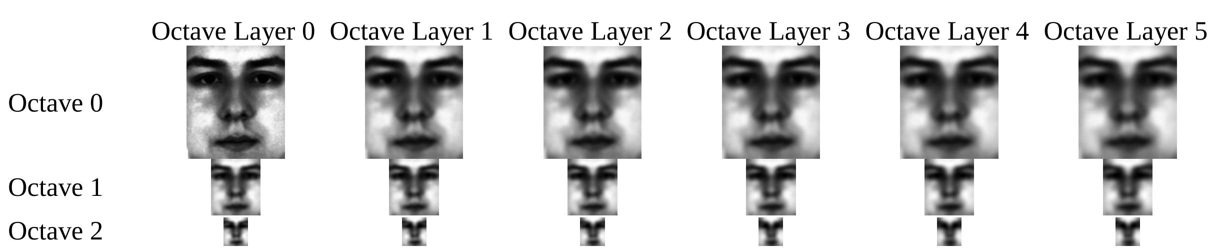

Fig. 2. A Gauss scale space with $o = 3$, $s = 6$ and $\sigma = 1{,}6$.

Gauss scale space that is generated by this method does not contain the input image. After the calculation of Gauss scale space for all octaves, the first image of the first octave is replaced with the input image to be able to obtain and compare the results provided in Table 1.

4. Face and Eyes Detection

Haar-like feature based cascade classifier is used to detect face and eyes to test the proposed face recognition framework. Implementation of face and eyes detection of the framework utilizes GPU by OpenCV library to increase the calculation performance (Bradski, 2000).

Cascade classifier consists of several simpler classifiers applied to a region of interest and evaluated together. First a classifier is trained with the same sized target object images of positive and negative examples. After training, the classifier is applied to the region of interest in an input image and it performs a prediction on the existence of an object in the region of interest. To detect an unknown sized object in an image, the classifier scans across the image at multiple scales and locations. Boosted cascade classifier is based on different boosting techniques. Haar-like features are used to train cascade classifiers (Viola and Jones, 2001), (Lienhart and Maydt, 2002). Feature prototypes of simple Haar-like and center surround features are shown in Fig. 3.

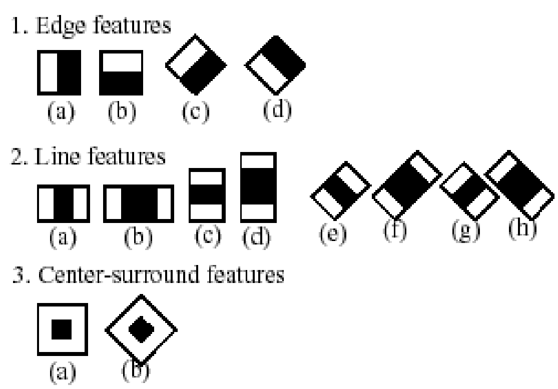

Fig. 3. Haar-like features used by cascade classifier. (Lienhart and Maydt, 2002)

5. Classification

After calculating the feature vectors, the next important step is the classification. In this study, *k*-Nearest Neighbor (*k*-NN) algorithm is used as the classification algorithm. *k*-NN classifier, which is the simplest classification algorithm, is based on the calculation of the distance between two samples. The class of the instance is assigned as the majority class of *k* neighbors. After several trials, NORM *L1* value is chosen as the distance metric and *k* equals 1.

6. Experiments and Results

The performance of LZM in low-resolution face images is evaluated using FERET database. Several tests are conducted with different settings and parameters. Then a face recognition framework is designed according to the test results.

6.1. Tests on Low Resolution and Blurred Images

All probe sets and gallery images are normalized to have zero mean and unit variance after they are cropped, and their sizes are fixed to 130×150. Then, Gauss scale spaces are calculated. After a set of tests, the scale space parameters are selected as $o = 3$, $s = 6$ and $\sigma = 1.6$. As a result, by using these parameters, scale spaces that contain 18 images ($Id = 0-17$) with resolutions 130×150, 65×75 and 32 ×37 from each normalized image are generated.

H-LZM feature vectors are calculated for each scale space image with the following parameters; the moment order $n = 4$, the kernel size $k = 5$, the grid size $N = 10$ and the number of histogram bins $b = 24$. The length of the feature vector with these parameters is 26064.

LZM in low-resolution face images are tested with several settings. First, classification ratios of each probe set scale space images $(Id = 0-17)$ with gallery scale space images $(Id = 0)$ are calculated. So, the calculated ratios for probe set scale images $(Id = 0)$ and gallery scale images $(Id = 0)$ give exactly the same ratio with the result provided in Table 1. On the other hand, it is noticed that the classification ratios for each probe set scale space image $(Id = 1-17)$ and gallery scale image $(Id = 0)$ decrease dramatically. These results are shown with legends H-LZM and H-LZM (Weighted) for four probe sets in Fig. 4.

For the second setting, three classification ratios for octaves between probe set and gallery scale space images are calculated: the first classifier matches probe set scale images with $(Id = 0-5)$ and gallery scale images with $(Id = 0-5)$, the second classifier matches probe set scale images with $(Id = 6-11)$ and gallery scale images with $(Id = 6-11)$, the third classifier matches probe set scale images with $(Id = 12-17)$ and gallery scale images with $(Id = 12-17)$. The results are given with legends H-LZM intra octave and H-LZM intra octave (Weighted) for four probe sets in Fig. 4.

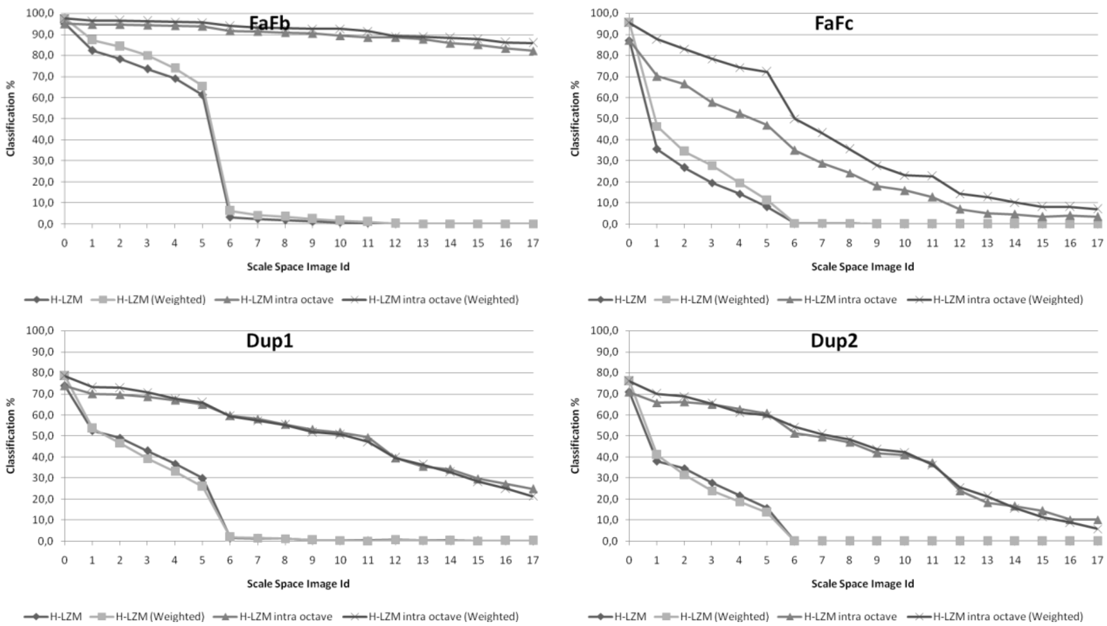

Fig. 4. The test results for low resolution and blurred images.

The two classification settings are examined. The first one matches probe set images with gallery images that are in different octaves. On the other hand, the second one matches probe set images with gallery images that are in the same octave. As a result, LZM gives better results for matching the probe set and gallery images that are in the same octave and even gives promising results at descending resolutions compared to the first setting.

6.2. Fusion of Octave Classifier's Results

The proposed face recognition framework's block diagram is given in Fig. 1. It takes an input image, detects face and eyes. Then, Gauss scale space is generated to utilize H-LZM feature vectors after the preprocessing step of the detected face and eyes image. So, H-LZM feature vectors of scale space images are ready for the classification part.

According to the above test results on FERET database, the classification part of the low-resolution face recognition framework is formed by fusing results of the octave classifiers. Three methods are developed to assess the test results which classify them according to class id majority, fusing class id majority of octave layers for each classifier or fusing class id of weighted octave layers for each weighted octave classifier. Among these, the best results are achieved with the last method.

An output of the proposed framework on real world video is shown by Fig. 5.

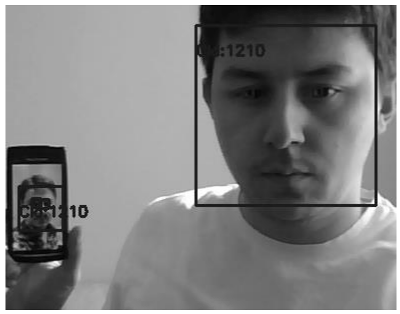

Fig. 5. An output of the proposed framework.

7. Conclusion

In this paper we have formulated a method that uses Local Zernike Moments (LZM) and Gauss scale space for face recognition in low-resolution face images. The performance of LZM in low-resolution face images is examined on FERET database. Then a face recognition framework is designed according to these test results. Results show that the proposed framework is promising for real world applications.

References

Biswas S., Aggarwal G., Flynn P.J. (2011). Pose-robust recognition of low-resolution face images. Computer Vision and Pattern Recognition, pp. 601–608.

Bradski, G.(2000). The OpenCV Library. Dr. Dobb’s Journal of Software Tools.

Hennings-Yeomans P.H., Baker S., Kumar B.V.K.V. (2008). Recognition of Low-Resolution Faces Using Multiple Still Images and Multiple Cameras. Biometrics: Theory, Applications and Systems, pp. 1–6.

Jafri R., Arabnia H.R. (2009). A Survey of Face Recognition Techniques. Journal of Information Processing Systems, vol.5, no.2, pp. 41–68.

Khontanzad A., Hong Y.H. (1990). Rotation invariant image recognition using features selected via a systematic method. Pattern Recogn, vol. 23, pp. 1089–1101.

Li B., Chang H., Shan S., Chen X. (2010). Low-Resolution Face Recognition via Coupled Locality Preserving Mappings. Signal Processing Letters, vol. 17, no. 1, pp. 20–23.

Lienhart R., Maydt J.(2002). An Extended Set of Haar-like Features for Rapid Object Detection. IEEE ICIP, vol. 1, pp. 900–903.

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

Rizvi S.A., Phillips P.J., Moon H. (1998). The feret verification testing protocol for face recognition algorithms. Computer Engineering, no. October, pp. 48.

Sariyanidi E., Dagli V., Tek S.C., Tunc B., Gokmen M.(2012). Local zernike moments: A new representation for face recognition. International Conference on Image Processing, pp. 585–588.

Viola P., Jones M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. Conference on Computer Vision and Pattern Recognition, pp. 511–518.

Zhou C., Zhang Z., Yi D., Lei Z., Li S.Z. (2011). Low-resolution face recognition via Simultaneous Discriminant Analysis. International Joint Conference on Biometrics, pp. 1–6.