Illumination-Robust Face Recognition from a Single Image per Person Using Matrix Polar Decomposition

Mehdi EZOJI†, Student Member and Karim FAEZ‡, Member

SUMMARY In this paper, a novel illumination invariant face recognition algorithm is proposed for face recognition. This algorithm is composed of two phases. In the first phase, we reduce the effect of illumination changes using a nonlinear mapping of image intensities. Then, we modify the distribution of the coefficients of wavelet transform in certain sub-bands. In this step, the recognition performance is more important than image quality. In the second phase, we used the unitary factor of polar decomposition of enhanced image as a feature vector. In the recognition phase, the correlation-based nearest neighbor rule is applied for the matching. We have performed some experiments on several databases and have evaluated the proposed method in different aspects. Experimental results in recognition show that this approach provides a suitable representation for overcoming illumination effects.

key words: face recognition, illumination invariant, matrix polar decomposition, wavelet domain, single image per person

1. Introduction

One of the major goals of computer vision and pattern recognition is illumination invariant recognition method. Illumination variation can strikingly alter the object appearance. One of the most relevant and important research fields is illumination invariant face recognition. During the past few decades, major advances have occurred to resolve this problem in the controlled environment and various approaches have been proposed. Current research efforts generally fall into three main categories:

- **Photometric Normalization**: in this approach, face images are preprocessed to normalize the images to appear stable under varying illumination. Histogram Equalization (HE) is the most commonly used approach even for images with controlled illumination [1]. Recently, block-based histogram equalization [2] and region-based histogram equalization [3] have been proposed. A local normalization technique which can effectively eliminate the effect of uneven illumination proposed in [2]. In this method, the local statistical properties of the processed image kept the same as that of image under normal illumination. In [3] illumination normalization was based on the HE, Gamma intensity correction and quotient image with a region based strategy to eliminate the side-lighting effects. Du et al. [4] performed illumination normalization in the wavelet domain. Chen et al. [5] employed a discrete cosine transform on the face images to compensate the illumination variation effects in the logarithm domain. Choi et al. [6] could find the shadow characteristics according to the direction of light and compensate it.

- **Illumination Invariant Feature Extraction**: Adini et al. [7] through an empirical study indicated that none of the conventional representations is sufficient by itself to overcome image variations because of the change in the direction of illumination. Gao and Leung [8] proposed Line Edge Map (LEM) for face recognition. The edge pixels are grouped into line segments and a revised Hausdorff distance is designed to measure the similarity. Zhao and Chellappa [9] presented an illumination insensitive method based on symmetric shape from shading under the assumption of a fixed and symmetric shape for real faces. A statistical shape from shading model to recover face shape using a single image and to synthesize the same face under new illumination developed in [10]. Based on the assumption that faces have the same shape but differ in the surface albedo, Shashua et al. [11], [12] proposed the Quotient Image (QI) method. Chen T. et al. [13], [14] modified QI, based on minimizing the total variation of image. Wang et al. [15] extend QI to estimate the illumination part using a simple Gaussian filter. Kanan et al. [16] proposed a novel face recognition approach based on Adaptively Weighted Patch Pseudo Zernike Moment Array (AWPPZM) when only one examplar image per person is available.

- **Face Modeling**: the face model is constructed based on a statistical or physical model. In statistical modeling such as Eigenface, Fisherface and Laplacianface [17], [18], no information about face properties is needed. In physical modeling, certain object reflectance properties such as lambertian reflectance are assumed. For lambertian surface in absence of the shadows, the surface shape and the albedo can be recovered from three images of this surface under known and independent light sources [12]. [19], [20] describe new methods that match the statistical models of the shape and texture of faces to a new image. Batur and Hayes [21] generalized the 3-D linear subspace model so that it is not affected by the regions in shadow. Belhumeur and Kriegman [22] proved that the set of images in fixed pose, but under different illumination conditions, form an illumination convex cone. Ignoring the cast shadow, intensity of a convex lambertian object can be approximated by a 9-D linear subspace based on spherical harmonic representation [23]. Lee et al. [24] showed that there exist configurations of $k$ point light source directions, with $k$ typically ranging from 5 to 9, such that by taking $k$ images of an ob-

Manuscript revised April 7, 2009.
†The authors are with Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran.
a) E-mail: m_ezoji@aut.ac.ir
DOI: 10.1587/transinf.E92.D.1561

Copyright © 2009 The Institute of Electronics, Information and Communication Engineers
ject under these single sources, the resulting subspace is an effective representation for recognition under a wide range of light conditions. In some approaches described before, the surface normal information is recovered from the 2-D intensity images. Recently several approaches such as [19] are designed to use the 3-D information acquired by active sensing devices like 3-D scanners or stereo vision systems.

However, in face modeling approaches, the assumption of convex lambertian surface (i.e. ignoring the cast shadow and specularity problem) is not valid for real faces. Besides, a number of images of the subject under varying illumination conditions or 3-D shape information are needed during the training phase.

A history of results on illumination invariant face recognition algorithms is provided in [25]. Nevertheless, the performance of most existing algorithms in their original form is highly sensitive to varying illumination condition in particular. As mentioned in [7], [25], it has been proven that the differences in the shading and shadows on the face appearance due to illumination and viewing direction can be larger than the inherent differences between individuals. None of the conventional invariant representations can overcome image variations [4], [26]. However, not only the performance of these methods are still not satisfactory but also many of them are based on the assumption that the shape of faces or the lighting modes of the images are known or can be estimated and rely on the volume of training set. It is worthwhile to highlight several distinct advantages of our method here:

- The proposed method doesn’t require the main features of face such as the shape and reflectance properties and illumination condition.
- Unlike many of the other techniques, the proposed method does not require a training set containing images of the same person under varying illumination condition.

To the best of our knowledge, this is the first work on face representation which considers the Matrix Polar Decomposition (MPD) of face images. The experimental results obtained on the most common database in this field, validate that our method has an acceptable performance in comparison with the other methods reported in the literature. The rest of the paper is organized as follows: first, the illumination normalization approach is described. Then a tolerant representation of face images in wavelet domain against illumination changes is presented. In Section 3, MPD is reviewed and the proposed method based on this decomposition is introduced. In Sect. 4, we provide experimental results on face recognition. Finally, conclusion is given in Sect. 5.

2. Preprocessing and Normalization in Wavelet Domain

In this phase, face images are processed and enhanced to diminish the effects of varying illumination conditions. Since, this phase usually influences the overall performance of the face recognition algorithm, it must be carefully designed. To do this, there are two main steps in our proposed method. Firstly, we normalized the images while keeping the main facial features unimpaired. Secondly, the wavelet components of the normalized image are modified to appear significantly stable under illumination changes. In this step, recognition performance is more important than image quality. The block diagram of the algorithm is illustrated in Fig. 1.

2.1 Preprocessing

There is physiological evidence that the response of the retina cells is nonlinear to the intensity of the incoming image. This nonlinear response usually is approximated as a log function of the intensity [7]. The nonlinear function which is proposed in this research is as follows:

\[
f(t) = \frac{\ln(1 + at^b)}{\ln(1 + a)}
\]

\(t\) is the intensity value of the input image at pixel \((x, y)\) and \(f(t)\) is the intensity of the normalized image. \(f(t)\) is monotonically increasing in the interval \([0,1]\). Consequently, it preserves the order from black to white in the gray scale. In addition, \(f(0) = 0\) and \(f(1) = 1\). A family of graphs with different \(a\) and \(b\) are shown in Fig. 2. The appropriate values of \(a\) and \(b\) are obtained empirically considering to the recognition performance and the corresponding curve with \(a = 0.01, b = 0.25\) is given in Fig. 2 as a Bold line.
2.2 Wavelet-Based Decomposition and Modification

After reduction of illumination variations by the method described in previous section, the wavelet based representation is used in this step. The main idea of this step is that firstly, illumination usually changes slowly [5], [15]. Consequently, illumination variations basically lie in the low frequency band. Secondly, according to cognitive psychological studies [28], [29] and experimental results [8], [30], edge map of face images are rich in information. This information contained the high frequency components of face image. Motivated by these facts, we want to extract multiple sub-band face images. Wavelet transform is the most popular multi-resolution analysis technique to break the image down into coarse and fine features. Each of the subspaces is obtained by taking the original incoming image and filtering it with a combination of high-pass and low-pass filters. These filters designed to maximize the amount of information obtained within each subspace. This decomposition can be repeated for n-levels. The image can later be reconstructed from these subspaces. In one hand, by removing or modifying low frequency sub-bands that contain the effects of illumination variation, we can significantly reduce these effects [5], [31]. On the other hand, as mentioned in [32], applying the histogram equalization improves the performance in face recognition.

In this research, we modify low frequency component and accentuate details coefficients. The underlying idea of this strategy is as follows: by performing the histogram equalization, the dynamic range of the pixel intensities is expanded. As a result, the dynamic range of the transformed coefficients will be expanded. According to this fact, as shown in Fig. 1, firstly we equalized the histogram of the coefficients in low frequency sub-band and modified the detail coefficients. Then, the histogram equalization is applied to all transformed coefficients, before reconstruction. The Daubechies wavelets, with their compact support and orthonormal nature, are one of the widely used wavelet families. Daubechies-1 wavelet with its efficiency in recognition performance in comparison with other wavelets is used in this research. Figure 3-d depicts the result of the proposed algorithm at the end of its first phase.

3. Matrix Polar Decomposition (MPD)

The key point of face recognition techniques is to develop appropriate representation for face space. Classical approaches such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are proposed for this representation. PCA is an unsupervised method which discovers the global Euclidean structure of the face space and LDA is a supervised method which seeks directions that are efficient for discrimination by maximizing the ratio of between-class variance to within class one [17]. Some extensions of PCA such as Modular PCA [33] has been studied and applied to face recognition. In modular PCA, a face image is partitioned into several sub-images and then a single conventional PCA is applied to each of them. Another approach for image analysis is the non-negative matrix factorization (NMF) which is based on the fact that image pixels are known to be non-negative [34]. Considering the illumination invariant face recognition, we find out that the unitary factor in the matrix polar form (defined later) of a face image can be regarded as an appropriate representation of the image. Experimental results in this research show that the MPD-based representation of a human face image contains rich identity information. The prime difference between these three methods and MPD is that MPD represents a face image without considering to other face images and dimensionality reduction. Also, PCA does more of feature classification and LDA does data classification while MPD does feature extraction.

For a full rank rectangular matrix \( M \in \mathbb{R}^{m \times n} \), the polar decomposition theorem asserts the existence of a factorization as follows:

\[
M = P_{m \times n} \times Q_{m \times n}
\]

where \( P \) is unitary and \( Q \) is Hermitian and positive factor.
In other word, \( P'P = I \) and \( Q = Q' \) [35]. In [36, 37] iterative and simple algorithms based on Newton’s method for the computation of the polar decomposition factors were described. In [36], orthogonal factor of \( M \) is computed efficiently by averaging the matrix and its inverse transpose in an iterative manner until convergence. Obviously, each column of Matrix \( P \) contains a basis vector while each column of \( Q \) contains the weights needed to approximate the corresponding column of \( M \).

**Remark 1:** Every SVD of \( M \) also defines the polar decomposition of it. Every matrix \( M \in \mathbb{R}^{m \times n} \) can be decomposed into \( M = U \Sigma V' \) where \( U \in \mathbb{R}^{m \times n} \) and \( V \in \mathbb{R}^{n \times n} \) are unitary and \( \Sigma \in \mathbb{R}^{m \times n} \) is a diagonal form. It can be shown that:

\[
M = \frac{U V^{'}}{P} \Sigma \frac{V^{'}}{Q}
\]  

(3)

subsequently, \( P = UIV' \) is the whitened version of \( M = U \Sigma V' \).

**Remark 2:** A mathematical theorem asserts that for every two real matrices \( A \) and \( B \in \mathbb{R}^{n \times n} \), there is an orthogonal matrix \( P \) such that for any orthogonal matrix \( X \):

\[
\| A - BP \|_F \leq \| A - BX \|_F
\]  

(4)

If SVD of \( B'A \) is \( U \Sigma V' \), then \( P = UV' \) holds this inequality [38].

**Corollary:** The unitary factor of polar decomposition of a matrix is the closest unitary matrix to it in the sense of Frobenius norm. That is, \( P \) solves the optimization problem which is as follows:

\[
P = \arg \min_{X} \| M - X \|_F
\]  

(5)

**Proof:** let \( B = I^{m \times n} \) and \( A = M \), in Eq. (4), the proof is immediate from aforementioned theorem [38] and Eq. (3).

Figure 4 shows the MPD-based unitary factor of a face image. As shown in Fig. 4, applying MPD reduces the effects of illumination variations. Motivated by this corollary, in this research, the unitary factor extracted from the enhanced face image is reshaped and regarded as a feature vector.

### 4. Experimental Results

In this section, we evaluate the performance of our proposed method by several experiments based on different databases and under comparison with several benchmark methods. Each of them attempt to deal with the recognition problem on different and important aspects. The databases used here include the Yale-a database, the Yale-B database [27] and the AR database [39]. These three face databases contain face images with variations in the lighting condition. In our experiments, all the images were cropped and normalized (in scale and orientation) such that the center of each eye is placed at a distance of 40 pixels from its closest borders and the distance between the two eyes is 80 pixels in an image of \( 160 \times 160 \). It should be mentioned that only face images without hair are considered in our research.

#### 4.1 Databases (DBs)

##### 4.1.1 Yale-B Database

The database contains images of 10 subjects each seen under 576 viewing conditions (9 poses \( \times \) 64 illumination conditions). These face images are divided into five subsets according to the lighting angle, in accordance with [40]. You can find a sample image per Subset in Fig. 5.

##### 4.1.2 Yale-a Database

The Yale-a Face Database contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, as shown in Fig. 6. The images demonstrate variations in illumination conditions, facial expressions and occlusions (with/without glasses).

##### 4.1.3 AR Database

This database contains over 4000 color images corresponding to 126 people’s faces (70 men and 56 women). Images
include frontal view faces with different illumination conditions, facial expressions and occlusions (sunglasses and scarf) which captured in two different sessions. However, some images were found lost or corrupted, there were 117 sets of usable images (64 men and 53 women). Since we are only concerned with illumination variation problem, the first four images of each subject under varying illumination are used in our experiments, as shown in Fig. 7.

4.2 Face Recognition

In this paper, we perform the closed-set evaluation, i.e., every probe is from the same group of subjects in gallery set. In general, there are \( n \) images per subject in the gallery set which are chosen randomly. For \( n = 1 \), one frontal image of each subject with normal illumination and neutral expression was used as a gallery sample. The distance metric which is used in this research is defined as follows:

\[
d(u, v) = \frac{1}{n} \sum_{i=1}^{n} \frac{|u_i - v_i|}{|u_i| + |v_i|}
\]

where \( u_i \) and \( v_i \) represent the \( i \)th element and \( |u| \) and \( |v| \) represent the absolute of vectors \( u \) and \( v \), respectively. We obtain a better performance when this distance measure is used instead of the Euclidean distance. Then, the nearest neighbor rule is used to classify the face images. Since some existing methods are applied on the same database, their reported results are directly listed from the corresponding papers.

4.2.1 Face Recognition under Controlled Condition

In this section, the AR database is used. Similar to [8], the neutral face images in the first session are used as the gallery set and the neutral face images in the second session are used as the probe set. The recognition results are summarized in Table 1. Eigenface uses PCA to project the data points along the directions of maximal variance which minimize the reconstruction error. Eigenface approach will work well as long as the test image is similar to the ensemble of images used in the calculation of eigenface [41] and the gallery set should include multiple images for each person with some variation. Because of these facts, PCA-based methods achieve low recognition rate when only one training image per person is used.

4.2.2 Face Recognition under “Leaving-One-Out” Strategy on the Yale-a Database

In this section, the error rate was evaluated using the

![Fig. 7 Illustration of 4 cropped images of one subject from AR database [39].](image)

Table 1

<table>
<thead>
<tr>
<th>Lighting Condition</th>
<th>Left light On</th>
<th>Right light On</th>
<th>Both light On</th>
<th>Using 3 Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>k=20</td>
<td>64.69</td>
<td>23.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>k=60</td>
<td>64.27</td>
<td>7.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>k=117</td>
<td>76.07</td>
<td>69.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>k=64 w/o 1st 3</td>
<td>48.72</td>
<td>10.26</td>
<td>55.4</td>
</tr>
<tr>
<td>Modular PCA</td>
<td>64</td>
<td>64.69</td>
<td>2.56</td>
<td>65.81</td>
</tr>
<tr>
<td>AWPPZMA [16]</td>
<td>76.36</td>
<td>64.69</td>
<td>42.74</td>
<td></td>
</tr>
<tr>
<td>The Proposed Method</td>
<td>97.53</td>
<td>96.30</td>
<td>97.15</td>
<td>96.85</td>
</tr>
</tbody>
</table>

“Leaving-One-Out” strategy [42] on the Yale-a database. In this strategy, the gallery set is constructed using all images of a subject except one, which must be classified. Then using the rest of images, the probe image is classified to the closest centroid of subjects. This process is repeated until all images are classified. Comparison results together with the results conducted by [8], [16], [34], [43] are shown in Table 2. From this table, it can be seen that the proposed method achieved the highest recognition rate.

4.2.3 Face Recognition under Varying Illumination

In this section, we design the experiments using the images taken under different lighting condition from AR database and Yale-B database as follows:

- **AR DB**: In this test, the face images in neutral expression with background illumination (the first image from left in Fig. 7) were regarded as a single model of each subject. The images under three different lighting conditions were used as probe images (the rest images in Fig. 7). The experimental results are tabulated in Table 3. We can see from this table that the proposed method significantly outperformed all the other methods. The variations of lighting conditions affected dramatically the performance of PCA-based approaches. According to these results, Belhumeur et
Table 4  Error rate under illumination variation on the Yale-B DB.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Subset 1</th>
<th>Subset 2</th>
<th>Subset 3</th>
<th>Subset 4</th>
<th>Subset 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Image</td>
<td>4.17</td>
<td>23.33</td>
<td>53.33</td>
<td>75.0</td>
<td></td>
</tr>
<tr>
<td>Histogram Eq.</td>
<td>0</td>
<td>5.85</td>
<td>36.0</td>
<td>33.0</td>
<td></td>
</tr>
<tr>
<td>Linear Subspace [40]</td>
<td>—</td>
<td>0</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCT-based [5]</td>
<td>—</td>
<td>0</td>
<td>0.18</td>
<td>1.71</td>
<td></td>
</tr>
<tr>
<td>Cone cast [40]</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TVQI [13]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LTV [14]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>The Proposed Method</td>
<td>0</td>
<td>0</td>
<td>1.43</td>
<td>1.05</td>
<td></td>
</tr>
</tbody>
</table>

Table 5  Recognition rate on Yale-B DB with different modification on detail coefficients.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Recognition Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = 1$</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>$\beta = 1$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$\gamma = 0.9$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>$\beta = 1$</td>
<td>$\gamma = 0.9$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>$\gamma = 0.9$</td>
</tr>
<tr>
<td>$\beta = 1$</td>
<td>$\gamma = 0.9$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>$\gamma = 0.9$</td>
</tr>
<tr>
<td>$\beta = 1$</td>
<td>$\gamma = 0.9$</td>
</tr>
</tbody>
</table>

al. [43] suggested that by discarding the first three principal components, variations due to lighting can be reduced. This case is denoted by ‘w/o 1st 3’ in Tables 2 and 3. However, there is no guarantee for this suggestion, because the eigenface method is highly dependent on the training samples but our experiment is based on a single model view. Besides, some useful identity information is weakened by discarding these components. Although the LEM [8] outperformed PCA-based method, when both lights were on, the recognition rates became much less than that of only one light on. However, our method is tolerant of extreme lighting conditions variations. Since, the database used in [8] was not exactly similar to our database in this experiment, their reported results are not given directly in Table 3.

- Yale-B DB: The Yale-B Database is commonly used to evaluate the performance of face recognition under varying illumination. To compare with other methods, similar to their experiment, the subset-1 is selected as the gallery set and other subsets are used as the probe set. The experimental results are tabulated in Table 4. Although the performance of our method is slightly less than [5] and [13], [14], the accuracy of the proposed method is still high and acceptable. It should be mentioned that the illumination cone method [40] needs seven face images per person to obtain the shape and albedo of face. In Sect. 2.2, the distribution of the coefficients of the wavelet decomposition was modified. Here, in addition to that modification, we rescale each detail coefficients in different level of decomposition with multiplying them by different scale factors. $\alpha$, $\beta$ and $\gamma$ are these factors for detail coefficients in level-1, level-2 and level-3, respectively. Subset-1 is selected as the gallery set. We achieved 100% recognition rate on subset-2 and subset-3. Experiment results on subset-4 and subset-5 are shown in Table 5.

4.2.4 Face Recognition under Varying Illumination from a Single Image per Person

To evaluate the performance of the proposed method from a single model per person, some other experiments were conducted on Yale-B database. In these experiments, we set the parameters as follows: $\alpha = \beta = 1$ and $\gamma = 0.9$ according to Table 5. For each subject, $n$ images from subset-1 selected randomly to form the gallery set. In one experiment, the gallery set contains 10 images (i.e. $n = 1$). We achieved 100% recognition rate on the rest of subset-1. For the other subsets, the results are drawn in Fig. 8. In rank $k$, the successfully recognized images are within $k$ nearest neighbors of the gallery image.

In other experiment, there are $n$ images of each subject in the gallery set. For each subject, these images were randomly selected from subset-1 (so $n \leq 7$). The probe set contains all of images in other subsets. Table 6 presents the average and standard deviation of recognition rates we measured on 10 gallery sets of size $10 \times n$ which formed randomly. According to the results of these experiments, the proposed method does not rely on the size and representative of the gallery set but many reported techniques such as methods based on PCA, LDA and illumination cone [40] require a training set containing multiple images per person with some variations in illumination conditions.
Table 7 The influence of each phase on the performances of the proposed method.

<table>
<thead>
<tr>
<th>Test</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>w/o the log-mapping of intensity</td>
<td>6.10</td>
</tr>
<tr>
<td>w/o wavelet-based modification</td>
<td>8.06</td>
</tr>
<tr>
<td>w/o MPD-based representation</td>
<td><strong>41.0</strong></td>
</tr>
<tr>
<td>Complete form of the proposed method</td>
<td>1.28</td>
</tr>
</tbody>
</table>

4.2.5 Algorithm Evaluation

To evaluate the influence of applying MPD-based representation of face images, we designed several experiments on Yale-a database [27].

- In the first experiment, in order to investigate the performance of each phase, we employ the proposed algorithm without that phase using leave-one-out strategy. The Error results are tabulated in Table 7. These results demonstrate the influence of each phase in overall performance of the algorithm. We can see from Table 7, the proposed method w/o MPD-based representation of face image is about 40% inferior in recognition rate to that when MPD is used. MPD-based representation of face images has a key role in the proposed method.

- We determined the potential of MPD-based representation for highlighting the discrimination information. According to LDA technique and Fisher criterion, the transformation $W$ which minimizes the objective function $J(W)$ is more efficient for discrimination.

$$J(W) = \text{trace}((W S_b W)^{-1}(W S_w W))$$  \hspace{1cm} (7)

where $S_b$ and $S_w$ are between and within class scatter matrices, respectively which are defined as:

$$S_b = \frac{1}{N} \sum_{i=1}^{N} N_i (m_i - \bar{m})(m_i - \bar{m})'$$  \hspace{1cm} (8)

$$S_w = \frac{1}{N} \sum_{i=1}^{N} \sum_{x \in C_i} (x - m_i)(x - m_i)’$$  \hspace{1cm} (9)

where $\bar{m}$ is the global centroid of images, $N$ denotes the number of face images, $N_i$ is the number of images in $i$th class and $s$ is the number of subjects. Using Yale-a database, we calculated the objective function before and after applying the MPD. The effect on this objective function when unitary factors of face images were used, caused 88.24 percent decreasing (improvement) in Fisher criterion (from 47.0945 down to 5.5391). After resizing the face images to $40 \times 40$ pixels, we calculated the Fisher criterion at the end of each phase. The results summarized in Table 8. For each face images, this sequence of points depicts a strictly decreasing curve which is desired. To comparison with LTV method [14], we repeat this process for LTV-based representation vs. the variations of $\lambda$ as shown in Fig. 9. Fisher criterion of LTV-based representation achieved 6.22 on average. It should be noted that the proposed method has some adjustable parameters, while LTV model requires only one parameter.

- We evaluated the performance of MPD-based representation of face images for minimizing the total variation. In TV-L1 model, large-scale factor $l$ of an image noted $I$, can be obtained by minimization a variational problem which forces the output to be more regular [14]:

$$f(u) = \int |\nabla u| + \lambda \| I - u \|_{L^1}$$  \hspace{1cm} (10)

Beginning from $I$, [37] iteratively computes the unitary factor $P$ of face image $I$ by averaging $I$ and its inverse transpose until convergence. In an empirical study, for all face images in available databases [27], [39], we compute $f(P_k)$ at each iteration indexed by $k$. It is observed that this sequence of points depicts a strictly decreasing curve for each face images. In addition, using Yale-B database, for all images of each subject, we compute the sequence of $f(P_k)$. Then the average of these sequences is computed for each subject. Figure 10 illustrates these strictly decreasing curves.

Encouraging experimental results demonstrate that unitary factor of a face image in polar form, has a noticeable effect on the illumination invariant face representation and recognition in single model databases.

4.3 Conclusion

In this paper, we have dealt with the illumination variation problem in face recognition. We have proposed a new robust face recognition algorithm invariant to the illumina-
tion changes in the frontal face images. In this method, firstly, the effects of illumination changes are reduced during the preprocessing and normalization phase. Then, using the MPD, we extract feature vectors which are rich in identity information for matching. Experimental results demonstrated that this approach is effective in illumination invariant face recognition. To the best of our knowledge, this is the first MPD-based work on face representation and recognition. For future works, we can focus on shadow compensation in wavelet domain, because in presence of shadow, the coefficients of higher sub-bands of transformed face are affected. In addition, we can research theoretically on the power of MPD-based representation for discrimination and recognition.

Acknowledgement

This research was supported by Iranian Telecommunication Research Center (ITRC).

References


Mehdi Ezoji received B.Sc. and M.Sc. degrees in Electronic Engineering from Sharif University of Technology in 2002 and Amirkabir University of Technology in 2005 respectively. Currently, he is a Ph.D. student at the Electrical Engineering Department of Amirkabir University of Technology, Iran. His research interests include computer vision, pattern recognition, image analysis and Biometric authentication. Email:m_ezoji@aut.ac.ir URL: http://ele.aut.ac.ir/mezoji

Karim Faez received his B.S. degree in Electrical Engineering from Tehran Polytechnic University as the first rank in June 1973, and his M.S. and Ph.D. degrees in Computer Science from University of California at Los Angeles (UCLA) in 1977 and 1980 respectively. Prof. Faez was with Iran Telecommunication Research Center (1981–1983) before joining Amirkabir University of Technology in Iran. He was the founder of the Computer Engineering Department of Amirkabir University in 1989 and he has served as the first chairman during April 1989-Sept. 1992. Professor Faez was the chairman of planning committee for Computer Engineering and Computer Science of Ministry of Science, research and Technology (during 1988–1996). His research interests are in Pattern Recognition, Image Processing, Neural Networks, Signal Processing, Farsi Handwritten Recognition, Earthquake Signal Processing, Fault Tolerant System Design, Computer Networks, and Hardware Design. He is a member of IEEE, and ACM. Email: kfaez@aut.ac.ir