Robust Face Recognition Under Illumination Changes and Pose Variations

Hamid Hassanpour\(^a\), Omid Kohansal\(^a\), Sekineh Asadi Amiri\(^b,\ast\)

\(^a\)Faculty of Computer Engineering & IT, Shahrood University of Technology, Shahrood, Iran
\(^b\)Faculty of Technology and Engineering, University of Mazandaran, Babolsar, Iran

\begin{abstract}
There are many applications for face recognition. Due to illumination changes, and pose variations of facial images, face recognition is often a challenging and a complicated process. In this paper, we propose an effective and robust face recognition method. Firstly, we select those areas from the face (such as eyes, nose, and mouth), which are more informative in face recognition. Then SIFT (Scale Invariant Feature Transform) descriptor is utilized for feature extraction from the selected areas. SIFT descriptor detects keypoints in the image and describes each keypoint with a feature vector with length 128. To speed up the proposed method, PCA (Principal Component Analysis) is applied on the SIFT feature vector to reduce the vector’s length. Finally, Kepenekci matching method is used to assess similarity between the images. The proposed method is evaluated on the ORL, Extended Yale B, and FEI databases. Results show considerable performance of the proposed face recognition method in comparison with several state-of-the-arts.

\end{abstract}

\section{Introduction}

Face recognition plays a major role in biometrics. The main positive aspect of face recognition methods in comparison with other biometrics identification approaches is the simple accessibility to facial images. As an example, in identification applications based on fingerprint processing, the finger should first be specially placed according to some restricted disciplines. The identification process is then carried out. Whereas, human identification using face recognition can be done in a wider range without paying attention to the camera.

In uncontrolled environments, issues such as image degradations, occlusions, illumination variations, facial pose variations, scale variations, poor images quality can influence the face recognition rate \([1]\). Facial occlusions may occur for several reasons. Some people wear sunglasses, scarf, and hat. Occlusions can also occur with medical masks, moustaches, beard, hairs covering the face, and make up. Thus, robustness to these problems is serious in face recognition systems. Figure \([1]\) shows some of these problems, which influence performance of face recognition systems.

A huge number of algorithms have been proposed for face recognition. These algorithms can be divided into two categories: holistic and feature-based approaches. Holistic approaches use the whole face region for feature extraction. These approaches have high efficiency on well aligned and equally illuminated images \([2]\). In uncontrolled conditions, face recognition rate seri-
Robust Face Recognition Under Illumination Changes and Pose Variations — H. Hassanpour, et al.

Figure 1. Examples of Different Conditions in Facial Images, Which Make Face Recognition Process a Challenging Task.

Figure 2. Showing Neighboring of a Pixel to Determine the Exterma Keypoints.

Figure 3. The Possible Interactions Among the Agents, Organizations, and Environment.

Figure 1. Examples of Different Conditions in Facial Images, Which Make Face Recognition Process a Challenging Task.

Figure 2. Showing Neighboring of a Pixel to Determine the Exterma Keypoints.

Figure 3. The Possible Interactions Among the Agents, Organizations, and Environment.

Figure 1. Examples of Different Conditions in Facial Images, Which Make Face Recognition Process a Challenging Task.

Figure 2. Showing Neighboring of a Pixel to Determine the Exterma Keypoints.

Figure 3. The Possible Interactions Among the Agents, Organizations, and Environment.

In this paper, a new method for face recognition is proposed, which is robust against illumination changes, and pose variations. Firstly, regions of interest in facial image such as the eyes, nose, and mouth are detected. Then, the Scale Invariant Feature Transform (SIFT) descriptor is extracted from these regions. SIFT descriptor detects keypoints from the image and specifies each keypoint with a feature vector of length 128. Following, the Principal Component Analysis (PCA) is applied on the SIFT feature vector to employ features with more discriminative ability for face recognition. Finally, similarity between the images is computed using the Kepenekci matching method.

The rest of this paper is structured as follows: SIFT descriptor is reviewed in Section 2. Section 3 reviews related works on face recognition. Section 4 describes the proposed approach whereas; Section 5 presents the performance evaluation of the proposed method. Finally, the conclusions are drawn in Section 6.

2 SIFT Descriptor

SIFT was first presented by David Lowe [3]. It identifies points of interest in the image and provides local descriptors which specify the neighborhood of these points [2]. The SIFT descriptor basically contains four steps: extrema detection, elimination of unstable keypoints, orientation assignment, and descriptor computation [4].

In the first step, the extrema are determined in the image filtered using the Difference of Gaussian (DoG) filter. The input image is gradually down-sampled

and filtered using the Gaussian filter. This process ensures the scale invariance of the technique. Figure 2 illustrates the implementation of the DoG filters at different scales.

To detect the extrema points, each pixel of the DoG filtered image is compared with its neighbors. Eight pixels in the same level and nine pixels in the two above and below levels are considered as its neighbors. If the pixel value is maximum or minimum compared to all of its neighboring pixels, it is considered to be a candidate keypoint. Figure 3 demonstrates the neighboring pixels for a pixel of the DoG filtered image. In this figure, neighboring pixels are shown as black rectangles.

In the second step, stability is calculated for each candidate keypoint. Keypoints with low contrast and keypoints along edges are eliminated.

In the third step, orientation is calculated for each remaining keypoints. The computation is done according to the gradient orientations in the neighboring pixel. Indeed, the values are weighted by the gradient magnitudes from the neighboring of the keypoint.

In the last step, a descriptor is provided for each keypoint. A 16 × 16 neighborhood is considered for each pixel, and gradient magnitudes and orientations are calculated for each of them. Their values are weighted by a Gaussian window. The 16 × 16 neighborhood
is divided into 16 sub-regions of size $4 \times 4$. The orientation histograms are created for each sub-region. Finally, a feature vector containing 128 ($16 \times 8$) values is created for each keypoint. In this paper SIFT descriptor is extracted from the facial image.

3 Related Works

Feature extraction is one of the most important steps in any face recognition system. Among the huge number of facial feature extraction techniques, holistic approaches utilize the whole image for face recognition, such as PCA [5], LDA [6], and ICA [7]. Geometric features, which are considered as local-based approaches, measure the distance between eyes, the width and length of the nose, and the mouth size [8]. These techniques are not inherently robust against variations in facial pose and expression variations. Indeed, any change in the facial pose or expression may result in different facial geometric features.

The Local Binary Patterns (LBP) [9], Gabor Wavelet [10], SIFT [11], and Speeded-Up Robust Features (SURF) [12], as a few, are among the major approaches developed for local-based face recognition. SIFT descriptor attracted a considerable attention due to their acceptable accuracy. Independent component analysis (ICA) [13] and SURF descriptors, named SIFT/SURF descriptor. This method has a weak efficiency for images with poor quality. Because, SURF descriptor extracts only a few keypoints from these images. Generally, SIFT-based methods can be used in real time applications, due to their acceptable accuracy.

The approach in [19] proposed the Local Gabor Fisher Classifier (LGFC) for face recognition. The approach in [20] utilizes a 2D Gabor Wavelet Transform (2DWT) and a 2D Hidden Markov Model (2DHMM) for face recognition. The approach in [21] uses both Gauss-Laguerre and Log-Gabor filters for feature extraction. After that, PCA is applied on the feature vector to reduce its length. Finally the reduced feature vector is utilized by the K-Nearest Neighbor (K-NN) for face recognition.

In [22] a face recognition method using SIFT is proposed. In this method, the $K$-means algorithm is applied on keypoints extracted using SIFT. In [23] a face recognition approach was proposed for robustness against illumination variations. First, the whole facial image is divided into high-lighted and non-highlighted regions. Then a pixel-level transformation in log space is afterwards constructed. The final step is to extend this chromaticity invariance to color space by taking into account the shadow edge detection.

In [24] some salient patches are extracted from the face, then multi-scale LBP features are extracted from patches. This feature is a high-dimensional vector. In [25], the Dual-Cross Patterns (DCP) of salient keypoints are used to describe the face. In [26] a new local facial descriptor, named as the Structural Binary Gradient Patterns (SBGP), for facial images is proposed. It measures relationships between local pixels in the image gradient domain and encodes them into a set of binary strings. The approach in [27] employs the kernel discriminates analysis for extracting features from input images. Furthermore, Support Vector Machine (SVM) and K-NN are utilized to classify the face image based on the extracted features.

4 Proposed Method

Figure 4 exhibits a flowchart of the proposed method. In this method, firstly, image illumination enhancement is done to deal with illumination variation problems. Then, some regions of the face such as eyes, nose, and mouth, which have the ability for discrimination of the individuals, are extracted from the face. Following, SIFT descriptor are applied on the face image. Then SIFT keypoints on salient regions are only considered. After that, PCA is applied to each SIFT-based feature vector to reduce its length. It should be mentioned that this process is done on all the training images and a test image. Finally, the Kepeneck matching method is utilized to assign the test image on one of the training images. Although, a lot of SIFT-based papers have been proposed for face recognition, considering the SIFT keypoints only on the salient regions of the face is the novelty of the proposed method.
4.1 Illumination Enhancement

SIFT descriptor is not robust against unbalanced image illumination. Hence, in the proposed method, image illumination enhancement is applied on the input image using the technique introduced in [28]. This method, named Local Normalization, enhances image illumination using the following equation:

\[
G(x, y) = \frac{F(x, y) - \mu_F(x, y)}{\sigma_F(x, y)}
\]

where \(F(x,y)\) is the original image, \(G(x,y)\) is the enhanced image, \(\mu_F(x,y)\) and \(\sigma_F(x,y)\) are the estimation of local mean and local variance of \(F(x,y)\) respectively.

Although, there are other approaches such as [29] [30] [31] to improve image illumination, these methods have a high computational complexity. Figure 5 indicates an example of image illumination enhancement using these four methods.

4.2 Extracting Salient Regions

Eyes, nose, and mouth are the most important parts of facial images for face recognition. We named these regions as salient regions or regions of interest. In real conditions, because of variation in hair and beard, ignoring these regions and considering the salient regions will improve the recognition rate. Also, using only the salient regions improves robustness of the proposed method against partial coverages such as using hat.

In this paper, same as [32], eight-points from facial images including the points around the eyes, the tip of the nose, points around the mouth, and the center of the face are extracted. These eight-points are shown in Figure 6. It should be mentioned that in some of the facial images the points on the tip of the nose and the point on the center of the face are overlapped with each other. The method in [32] uses DPM-based classification to extract the mentioned eight-points. Indeed, a facial image is the input of the classifier and its output is the estimation of the above eight-points. Finally, using these points, effective areas are identified for face recognition (Figure 7).

4.3 SIFT Descriptor and PCA

In this paper, SIFT descriptor is applied on the facial image. SIFT descriptor extracts some keypoints and describes each keypoint with a feature vector of length 128. In this step, only the SIFT keypoints extracted from regions of interest are considered, and the rest are discarded. The SIFT descriptor specifies each keypoint using a feature vector with 128 values. It seems that all of the 128 values may not be effective to describe a point; hence, feature reduction can be applied on these feature vectors. This reduction increases the time efficiency. In the proposed method,
PCA is used for feature reduction. PCA produces an optimal linear transformation from the original data space to an orthogonal eigenspace in which the reduced dimensionality leads to a least mean squared reconstruction error [9].

4.4 Kepenekci Matching

The proposed method is applied on all of the training images and the test image. Then, the similarity between the test face feature vector and each of the training image feature vectors is computed using Kepenekci matching [18].

Kepenekci matching is a weighted combination of two measures. Let us call T a test image and G a training image. For each feature vector t from the face T, a set of relevant vectors g from the face G is determined. Vector g is relevant if following equation is satisfied:

$$\sqrt{(x_t - y_g)^2 + (y_t - y_g)^2} < \text{distanceThreshold}$$  (2)

where x and y are coordinates of the feature vector points.

The overall similarity of two faces is calculated according to the following equation:

$$\text{Sim}_{1T,G} = \text{mean}\{S(t,g), t \in T, g \in G\}$$  (3)

$S(t,g)$ represents the similarity of two feature vectors according to cosine measure. Generally, the above equation computes the average of similarities between each pair of corresponding vectors. Then, the face with the most similar vector to each of the test face vectors is specified.

The following equation indicates the second measure which is used in Kepenekci matching:

$$\text{Sim}_{2T,G} = \frac{C_G}{N_G}$$  (4)

where $C_G$ is the number of related points in the training face G with the test face T, and $N_G$ is the total number of feature vectors in G.

Kepenekci similarity measure is a weighted sum of these two similarities:

$$\text{Sim}_{F,T,G} = \alpha \text{Sim}_{1T,G} + \beta \text{Sim}_{2T,G}$$  (5)

The recognized face is the one with the highest similarity value.

According to the Kepenekci matching, recognition rate can be calculated for each facial face recognition method. Recognition rate is defined as the percentage ratio of correctly identified images to the total number of images.

![Figure 8](image.png) Example of One Individual in Different Illumination Conditions From The Extended Yale B Database.

5 Experimental Results

In order to evaluate the proposed method, ORL database [33] containing facial images with different viewing directions and poses; FEI database [34] containing facial images with different viewing directions; and Extended Yale B [35] containing facial images with different illumination conditions are used. In the following, experimental results on these three databases are evaluated.

5.1 Extended Yale B Face Database

For evaluating the proposed method in different illumination conditions, we tested our approach on the Extended Yale B database. This database contains 38 distinct individuals and 64 pictures for each individual. All of the images in this database were categorized into five different sets in terms of illumination condition. The first category contains images with the appropriate illumination and gradually the illumination is dimmed in the following subsequent sets. Figure 8 shows these five categories.

The images of the first category are considered as the training set and categories 3, 4, and 5 are considered for the test sets. Table 1 shows the recognition rate before and after image illumination enhancement. The results show that image illumination enhancement increases the recognition rate.

5.2 FEI Face Database

As mentioned in Section 5, merely considering SIFT features on the region of interest improves the face recognition rate and leads the robustness to facial pose variations. We tested our approach on the FEI
Table 1. The Recognition Rate Before and After Images Illumination Enhancement.

<table>
<thead>
<tr>
<th>Test Category</th>
<th>Before preprocessing</th>
<th>After preprocessing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 3</td>
<td>87.1</td>
<td>90.2</td>
</tr>
<tr>
<td>Category 4</td>
<td>47.5</td>
<td>58.7</td>
</tr>
<tr>
<td>Category 5</td>
<td>45.2</td>
<td>57.2</td>
</tr>
</tbody>
</table>

Figure 9. An Example of One Individual, From FEI Face Database, in Five Different Viewing Directions.

The FEI face database. This database contains 200 distinct individuals and 14 pictures (with different viewing directions) for each individual. Figure 9 shows an example of one individual in five different viewing directions.

First, one image of each individual was considered for the training stage and five images were used in the test stage. Following, three images of each individual were used in the training stage and five images were used in the test stage. These experiments were done on 50, 100, and 200 distinct individuals. Table 2 shows the effect of the salient regions on the face recognition rate. It can be observed that merely considering the SIFT descriptor on the salient regions of the face has a great influence on the face recognition performance.

As mentioned before, the SIFT descriptor addresses each keypoint with a feature vector of length 128. To speed up the proposed method, PCA has been applied on these feature vectors to reduce their length. Table 3 shows the face recognition rate with different lengths of the SIFT feature vectors. Overall, in the case of using PCA, the proposed algorithm works faster, however, to obtain more accurate results, PCA is not helpful.

5.3 ORL Face Database

The ORL face database is used to evaluate the robustness of the proposed method against different viewing directions and poses, as well as to compare the proposed method with several available methods. This database contains 40 distinct individuals and 10 pictures for each individual. Figure 10 indicates an example of one individual with ten different viewing directions and poses in the ORL database. Five images were considered as the training images and the remaining five images were utilized as the test images. The results are shown in Table 4. Results show that the proposed method has a higher recognition rate compared to the state-of-the-arts.

5.4 Time Complexity

Our proposed face recognition method was implemented using C# programing language. The experiments were run on a computer with a 2.2 GHz Intel Corei7 and RAM of 6 GB. The computational times for the proposed face recognition technique on the images with size $70 \times 70$ and the length of the feature vector 128 are shown in Table 5. In this table the number of extracted SIFT keypoints from the whole and salient regions of the face are also shown. The results indicate that the proposed method has very low execution time.

6 Conclusions

In this article, an effective method for face recognition was proposed. First, salient regions of facial image are extracted, then SIFT descriptor, and in the following, PCA are utilized for feature extraction. Finally, the Kepekeci matching method is used to compute the similarity between the images. The proposed method is robust against illumination changes, partial coverage, and pose variations. Also, to further enrich robustness of the proposed method against illumination variations, a preprocessing is done on the image. The proposed method was evaluated on the ORL, Extended Yale B, and FEI databases. The results showed that the proposed method outperforms the existing face recognition in terms of recognition rate.

References


Table 2. The Effect of the Salient Region on the Face Recognition Rate.

<table>
<thead>
<tr>
<th>Number of individuals</th>
<th>Number of picture for each individual</th>
<th>Whole face region</th>
<th>Salient region of face</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>One picture</td>
<td>Three pictures</td>
</tr>
<tr>
<td>50</td>
<td>83.8</td>
<td>99.6</td>
<td>90.4</td>
</tr>
<tr>
<td>100</td>
<td>82.5</td>
<td>99</td>
<td>89.8</td>
</tr>
<tr>
<td>200</td>
<td>80.6</td>
<td>98.2</td>
<td>87.9</td>
</tr>
</tbody>
</table>

Table 3. The Face Recognition Rate With Different Lengths of the SIFT Feature Vectors.

<table>
<thead>
<tr>
<th>Length of the feature vector</th>
<th>Number of individuals</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>90.4</td>
<td>89.8</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>89.8</td>
<td>89.1</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>88.1</td>
<td>88.1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>84.2</td>
<td>83.6</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Comparison of Different Methods With the Proposed Method in Face Recognition Rate on the ORL Database.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Training images number</th>
<th>Five images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gauss-Laguerre [21]</td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>Local Gabor [19]</td>
<td></td>
<td>92.5</td>
</tr>
<tr>
<td>SIFT-Kepenekci [18]</td>
<td></td>
<td>97.9</td>
</tr>
<tr>
<td>2D Gabor [20]</td>
<td></td>
<td>99</td>
</tr>
<tr>
<td>kernel discriminates analysis [27]</td>
<td></td>
<td>96</td>
</tr>
<tr>
<td>Proposed Method</td>
<td></td>
<td>99.5</td>
</tr>
</tbody>
</table>

Table 5. Computation Time of the Proposed Face Recognition.

<table>
<thead>
<tr>
<th>Number of training images</th>
<th>Number of extracted keypoints on the whole face region</th>
<th>Number of extracted keypoints on the salient region</th>
<th>Computational time (Milli Second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>55</td>
<td>30</td>
<td>160</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>19</td>
<td>57</td>
</tr>
<tr>
<td>20</td>
<td>53</td>
<td>29</td>
<td>252</td>
</tr>
<tr>
<td>20</td>
<td>31</td>
<td>19</td>
<td>126</td>
</tr>
</tbody>
</table>


[6] P.N. Belhumeur, J.P. Hespanha, and D.J. Krieg-


Hamid Hassanpour received the B.S. degree in Computer Engineering from Iran University of Science & Technology, Tehran, Iran, in 1994, the M.S. degree in Computer Engineering from Amirkabir University of Technology, Tehran, Iran, in 1997, and the Ph.D. degree in signal processing from Queensland University of Technology, Brisbane, Australia. He is currently a professor at Department of Computer Engineering, Shahrood University, Iran.

Omid Kohansal received the B.S. degree in Computer Engineering from Hakim Sabzevari University, Sabzevar, Iran, in 2013 and the M.S. degree in Computer Engineering from Shahrood University of Technology, Shahrood, Iran, in 2015.

Sekineh Asadi Amiri received the B.S. degree in Computer Engineering from Babol University, Babol, Iran, in 2006, the M.S. degree in Computer Engineering from Shahrood University of Technology, Shahrood, Iran, in 2010, and the Ph.D. degree in Computer Engineering from Shahrood University of Technology in 2016. She is currently an assistant professor at Department of Engineering, University of Mazandaran, Iran.