Disguised Face Recognition by Using Local Phase Quantization and Singular Value Decomposition

Fatemeh Jafari [a], Hamidreza Rashidy Kanan [b,*]

[a] Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
[b] Department of Electrical, Biomedical and Mechatronic Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

Received 11 November 2015; accepted 23 December 2015

Abstract

Disguised face recognition is a major challenge in the field of face recognition which has been taken less attention. Therefore, in this paper a disguised face recognition algorithm based on Local Phase Quantization (LPQ) method and Singular Value Decomposition (SVD) is presented which deals with two main challenges. The first challenge is when an individual intentionally alters the appearance by using disguise accessories, and the second one is when gallery images are limited for recognition. LPQ has been used for extraction of the statistical feature of the phase in windows with different sizes for each pixel of the image. SVD is used to cope with the challenge of the gallery images limitation and also with the help of original images extracted from that, every single image turns to three renovated images. In this study, disguise is intended as a blur in the image and Local phase quantization method is robust against the disguised mode, due to the use of the statistical feature of the Fourier transform phase. Also the use of different-sized window instead of fixed window in feature extraction stage, the performance of the proposed method has increased. The distance of images from each other is computed by using Manhattan and Euclidean distance for recognition in the proposed method. The Performance of the proposed algorithm has been evaluated by using three series of experiments on two real and synthesized databases. The first test has been performed by evaluating all the possible combinations of the different-sized windows created in the feature extraction stage, and the second experiment has been done by reducing the number of gallery images and then the third one has been carried out in different disguise. In all cases, the proposed method is competitive with to several existing well-known algorithms and when there is only an image of the person it even outperforms them.

Keywords: Disguised Face Recognition, Local Phase Quantization, singular value decomposition, Fourier Transform, Manhattan and Euclidean Distance.

1. Introduction

Face recognition is one of the most important issues in the field of pattern recognition and computer vision. The aim of face recognition is identification and authentication of people regards to their images or videos by using a database of their faces. Because of having some advantages e.g. high accuracy and low permeability, face recognition technology is used in information security, law monitoring, intelligent cards, access control and etc... In spite of the researches carried out in the past decades, face recognition is still dealing with some challenges due to the natural complexity of the problem. These challenges are caused by human face variations which

* Corresponding author. Email: h.rashidykanan@qiau.ac.ir

come from variation in illumination, covering and disguise [1].

Researchers are faced with a kind of challenge lately and it's when a person forges his identity deliberately; so the ability of accurate identification of a forged identity and with the minimum number of images from a person is getting important [2]. Many approaches are proposed to deal with several challenges such as artificial disguise, gesture and illumination and techniques e.g. Geometric features [3], Eigen Face [4], Local features [5], Neural networks [6], Elastic bunch graph matching [7], and wavelet [8].

Ramanathan et al. [9] used facial similarity for several variations, including disguise by forming two Eigen-space from two halves of the face which in one of them the left half and the other one right half of the face have been used. In the recent study by Singh et al. [2] a face recognition algorithm was proposed in order to overcome two main challenges. The first challenge is when a person intentionally changes his appearance by disguise accessories and the second challenge is when gallery images for recognition are limited. This algorithm uses dynamic neural network architecture for extracting fuzzy features of texture by using Gabor 2-D polar logarithmic transformation [2].

In disguised face recognition system, there is a disguised image that its true identity should be recognized. There are many of methods of impersonation that a person may use them in order to change his or her face. Each of these changes can alone produce much uncertainty that may cause many changes in pixel values of the image of a face. Since the type and the extent of these changes are not known in advance, could make the feature extraction in these systems to be faced with some problems. To cope with this challenge, in the feature extraction stage of the proposed method, for each pixel, windows of various sizes have been introduced instead of using a fixed-sized window. For example, if may assume a pixel around the eye, different dimensions of the window can lead pixel value toward the forehead or cheek areas, the areas that are likely to be less affected by changes. Another challenge in this field would be the limited area of the gallery in order to deal with that through the proposed method by using the singular value decomposition (SVD) and then from any individual image, three renovated image is being produced.

Section 2 presents the proposed method. Afterwards, the numerical result will be presented in section 3, where there are three kinds of experiments; the first one is review evaluate all possible combinations of different-sized windows which are created in the feature extraction stage and two others may evaluate the method with variety of changes in the face and to reduce the number of gallery images. At the end, conclusion is presented.

2. The Proposed Method

The proposed algorithm is based on local phase quantization (LPQ) method and singular value decomposition (SVD), which deals with two main challenges. The first is when an individual intentionally alters the appearance by using disguise accessories, and the second is when gallery images are limited for recognition.The proposed algorithm includes two parts: training and testing. In the training phase, at the first time SVD method has been applied to the gallery images and then three renovated images have been produced from each of them that through which number of training samples could be increased. Then by using LPQ descriptor for each image in the gallery, several features vector with different windows extracted and stored. Then, the recognition accuracy for all possible combinations of these windows is computed and Finally, Brute force algorithm is applied to these combinations and a combination with the highest recognition accuracy as the optimal combination is selected. In the test part, first the features are extracted from the image, then the distance to all extracted features in the previous part is calculated and the closest match is obtained to find the owner of the image. The overall structure of the proposed method is presented in figure 1. In the proposed method, original action would be the use of windows with different dimensions instead of the fixed window, and in order to increase the number of training samples by using the SVD method which may greatly improve the method performance. Experimental results confirm the issue. In the recognition level, the Euclidean and Manhattan distance are used because their computation is very fast and it causes high speed in recognition and makes the real time applications possible.

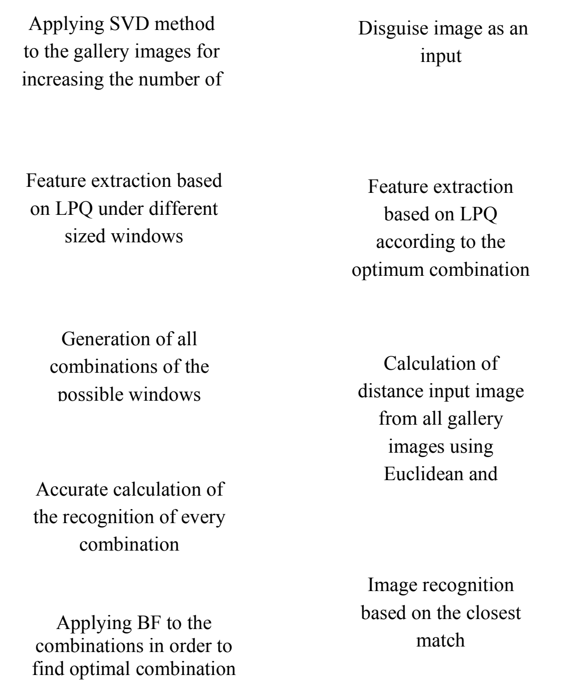

Fig. 1. Structure of the proposed method, a: training phase, b: testing phase

2.1. Singular value decomposition (SVD) of gallery image

To overcome with the challenge of the gallery images limitation (when only there is one image of per person); a solution is being considered based on the singular value decomposition method. By using this method, each face image is being decomposed to the original one and then three renovated images are produced from each of them. So that, if f(x) would be the image of the face, equation (1) represents the singular value decomposition method [10, 9] of that image.

$$f(x) = \sum_{i=1}^{n} \sigma_i . u_i . v_i^T \qquad (1)$$

Where u_i and v_i are the i-th column of u and v, respectively. u and v are the more specific vectors of ff^T and f^Tf. σ_i is the image f(x) singular values, so that $\sigma_2 \geq \sigma_1 \ldots\ldots \geq \sigma_n$ [10 , 9].

In the equation (1) image f(x) is produced from a set of original images i=1, n و $F_i = \sigma_i . u_i . v_i^T$. Figure 2 shows the original image and renovated images by using the original images based on the singular value decomposition. In this form, images have been produced according to the equation (1) and i is the value of one to three images. As seen, with increasing the value of n, renovated images get close to the original image. Experimental results have showed that if $n \geq 4$, the difference between the original image and renovated images would be very low.

Accordingly, from the single gallery that contains only a single training sample per person, a new gallery containing four pictures per person (one original and three renovated ones) could be produced. In the proposed method through this way, the challenge of single gallery image is partially resolved.

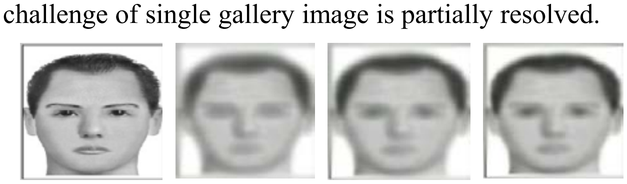

Fig. 2 . Renovated images by using SVD method with values i = 1 to 3

2.2. Feature Extraction Based on Local Phase Quantization (LPQ)

LPQ descriptor which proposed by Ojansivu and Heikkila has gained a lot of popularity in recent years due to its outstanding performance in image texture analysis [10]. LPQ method is based on quantized phases of the Discrete Fourier Transform (DFT). This descriptor utilizes phase statistical feature computed locally in a window for each pixel from the image, generating a code and finally making a histogram from the codes. The codes produced by the LPQ are insensitive to centrally symmetric blur, which includes motion, atmospheric turbulence blur and etc. [10]. Due to the images limited size, in practice, it is impossible to achieve absolute stability against the blur in images. Convolution of an ideal image with spread point blur function is much more than what we see in images therefore, some information is lost when the blur boundaries is much larger than the image size.

2.2.1. Blur Invariance Using Fourier Transform Phase in LPQ

In digital image processing, the discrete model for spatially invariant blurring of an original image $f(x)$ resulting in an observed image $g(x)$ that given by [10]:

$$g(x) = (f * h)(x) \quad (2)$$

Where $h(x)$ is the point spread function (PSF), $*$ denotes 2-D convolution and x is a vector of coordinates $[x,\ y]^T$. In the Fourier domain, this corresponds to:

$$G(\mathrm{u}) = \mathrm{F(u).H(u)} \quad (3)$$

Where $h(x)$ is the point spread function (PSF), $*$ denotes 2-D convolution and x is a vector of coordinates $[x,\ y]^T$. In the Fourier domain, this corresponds to:

$$|G(u)| = |F(u)||H(u)| \quad (4)$$

$$\angle G(u) = \angle F(u) + \angle H(u) \quad (5)$$

If it is assume that the PSF $h(x)$ is centrally symmetric, namely $h(x) = h(-x)$, its Fourier transform is real-valued, and as a consequence its phase is only a two-valued function, given by:

$$\angle H\ (u) = \begin{cases} 0 & if\ H(u) \ge 0 \\ \pi & if\ H(\mathrm{u}) < 0 \end{cases} \quad (6)$$

This means that:

$$\angle G(u) = \angle F(u)\ \ for\ all\ \ H(u) \ge 0 \quad (7)$$

As a result, the phase of the observed image $\angle G(\mathrm{u})$ at the frequencies, where $H(\mathrm{u})$ is positive, is invariant to centrally symmetric blur.

2.2.2. Short Term Fourier Transform (STFT)

Local phase quantization method is based on stability property of Fourier phase spectrum against blur which is determined by extracting local phase features by using 2-D DFT or more precise, calculating short term Fourier transformation (STFT) in a neighborhood:

$$F(u,x) = \sum_{y \in N_x} f(x-y) e^{-j2\pi u^T y} = w_u^T f_x \quad (8)$$

Where N_x is the neighbourhood and f(x-y) is the function value in the neighbourhood. A more detailed description in [10] can be found.

2.2.3. Statistical Analysis of the Coefficients and Quantization

It is assumed that the image function f(x) is the result of a first order Markov process where the correlation between adjacent pixels is ρ and variance of each sample is σ^2, covariance between pixels will be:

$$\sigma_{ij} = \rho^{\|x_i - x_j\|} \quad (9)$$

Where ||.|| is the L2 norm. Covariance matrix of all samples in $\mathcal{N}_x$ can be described as follows:

$$C = \begin{bmatrix} 1 & \sigma_{12} & . & . & . & \sigma_{1M} \\ \sigma_{21} & 1 & . & . & . & \sigma_{2M} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \sigma_{M1} & \sigma_{M2} & . & . & . & 1 \end{bmatrix} \quad (10)$$

Covariance matrix of transformation coefficients vector Fx can be obtained from the following equation:

$$D = WCW^T \tag{11}$$

It can easily be found that for ρ>0, D is not a diagonal matrix, which means that the coefficients are correlated.

First, the coefficients should be decorrelated, because it can be demonstrated that if the quantized samples are statistically independent, the information in scalar quantization maximally will be maintained. Therefore, Whitening Transform is used:

$$G_x = V^T F_x \tag{12}$$

Where V is an orthonormal matrix extracted from SVD of matrix D:

$$D = U \Sigma V^T \tag{13}$$

In the next level, G_x is computed for all positions in image and obtained vectors are quantized by using a scalar quantize:

$$q_j = \begin{cases} 1, & if \quad g_j \geq 0 \\ 0, & otherwise \end{cases} \tag{14}$$

The coefficients are quantized in integer values between 0-255 by using a binary coding as follows:

$$b = \sum_{j=1}^{8} q_j 2^{j-1} \tag{15}$$

Finally, a histogram is formed from these integer values and used as a feature vector of 256 elements for classification.

2.2.4. Local Phase Quantization in Face Recognition under Disguise Condition

In this study the disguised face image will be considered as an original image plus disguise as a blur function added to the image function and we try to find most similar original image function while have no precise information about disguise. Fig.1 shows quantized phase features obtained by different LPQ local window.

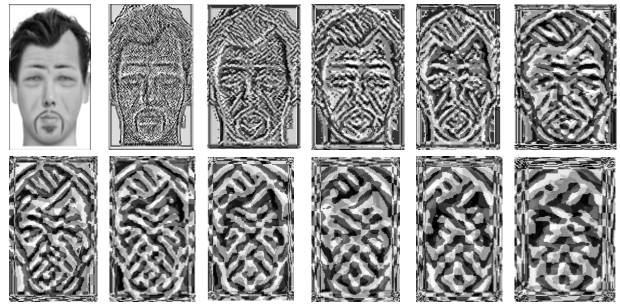

Fig. 3. Disguised face image and the corresponding quantized phase features obtained by different LPQ local window scale (small windows to large windows from left to right)

2.3.Recognition in the Proposed Method

After extracting features from images by using different windows, we'll have an M×N feature matrix for each image, where M is the number of windows and N is the length of the histogram (256). To calculate the distance an image from another image, The Euclidean distance and Manhattan distance are used which they are as follows:

Euclidean Distance:

$$d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + ... + (q_n - p_n)^2} \tag{16}$$

Manhattan Distance:

$$d(q,p) = |q_1 - p_1| + |q_2 - p_2| + ... + |q_n - p_n| \tag{17}$$

Where q_i is the i-th feature from image q ?

3. Experimental Results

In this section, first two databases to evaluate the performance of the proposed method will be described, then two different examinations on these databases are performed and the proposed method performance is compared with PCA [4], GF [3], LFA [5], IGF [11], LBP [12] and 2D-LPGNN [2].

3.1.Data Base

The proposed method is evaluated by using two databases. The first database is the AR face database and second is a database including synthetic disguised images that we were produced by using Faces software [14].

3.1.1. AR Database

The AR database includes about four thousand images from 126 person (70 men and 56 women). There are twenty six colored images in front view of each person with changes in facial expressions, glasses and scarves in two different sessions. In addition, there are illumination varieties in some images [15]. Fig. 2 shows representative images of an individual from the AR database. The AR database is divided into three categories:

- Expressions with illumination (14 images for each person)
- Glasses with illumination (6 images for each person)
- Scarf with illumination (6 images for each person)

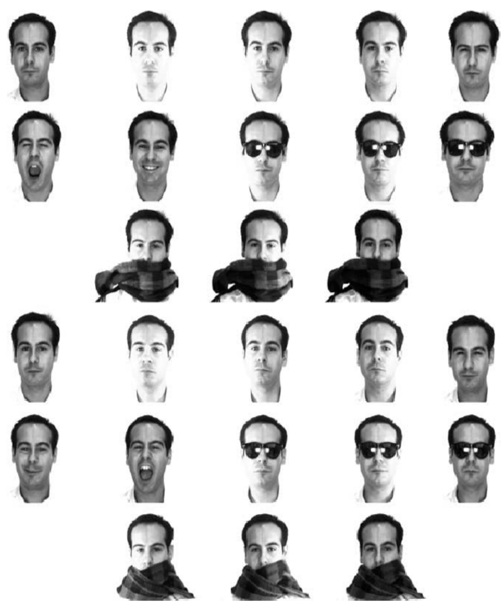

Fig. 4. Images of a same person in AR database [15].

3.1.2. Synthetic Disguised Data base

In this study, a database of faces by using Faces software [14] also has been prepared. The Synthetic Disguised database includes three thousand five hundred and seventy images from 85 person in 42 different modes (65 men and 20 women) along with the overall set of variations in the face. Fig. 3 shows representative images of an individual from the Disguised database. This database is used to study the performance of the proposed algorithm in different types of the disguise. The Synthetic Disguised database is divided into seven categories:

- Changes in hair style (6 images for each person)
- Changes in beard and moustache (6 images for each person)
- Changes in glasses (6 images for each person)
- Changes in hat (6 images for each person)
- Changes in lip / eyebrow / nose (6 images for each person)
- Changes in aging/wrinkles (6 images for each person)
- Combination of multiple changes (6 images for each person)

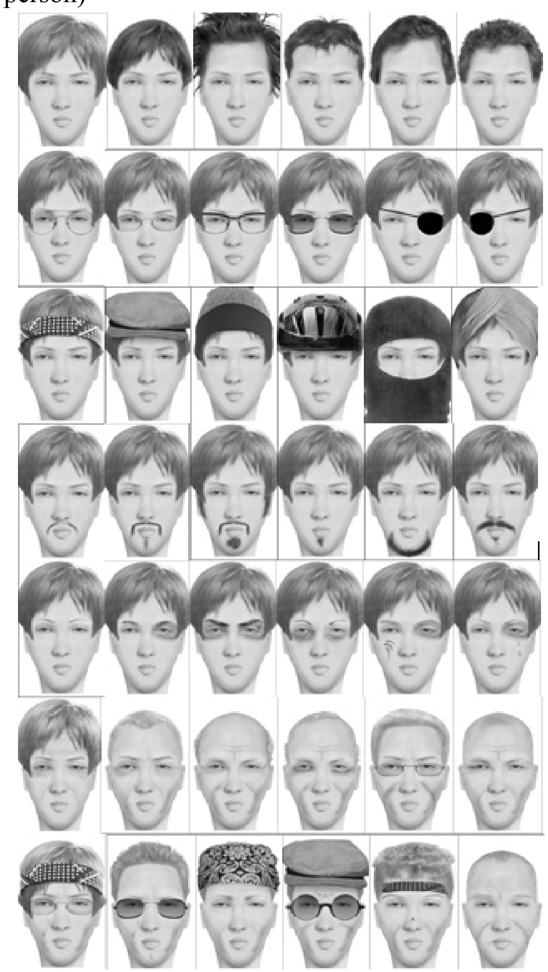

Fig. 5. Images of a same person in Synthetic Disguise Database

3.2. Selecting the Optimal Window in LPQ by Using Brute Force Algorithm

In LPQ feature extraction method, selecting the optimal window size has a significant effect on the performance of the proposed algorithm. For this purpose, several feature vectors were extracted from databases images with different window sizes. Resulting of the algorithm is shown in Table 1 and “Fig. 6” by using various windows:

Table 1

Performance of the proposed method on synthetic disguised database for different window sizes

window sizes (Pixels× Pixels)	Identification Accuracy (%)	
	Manhattan	*Euclidean*
3×3	51.05	32.35
5×5	59.38	41.04
7×7	64.22	50.00
9×9	67.93	58.12
11×11	67.44	57.7
13×13	66.53	58.68
15×15	69.33	62.25
17×17	67.09	60.99
19×19	66.18	57.14
21×21	66.88	58.54
23×23	65.27	57.98
25×25	66.04	57.91
27×27	63.17	55.18
29×29	61.76	56.79
31×31	63.38	55.88

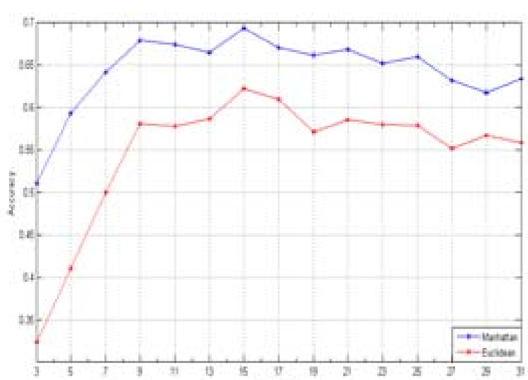

Fig. 6. Performance of the proposed method for different window sizes

As can be seen the Manhattan distance leads to a better result than the Euclidean distance in this case. To reach higher accuracy, the distance of the images is obtained by more than one window size and after averaging the distances the identification accuracy is calculated. For obtaining the best combination of windows, Brute Force algorithm is applied to all possible combinations and the one with the highest accuracy is selected as the global optimal combination. The identification accuracy by using the Manhattan distance and optimal combinations for disguise types are shown in the table 2.

Table 2

Obtained accuracy by using optimal combinations and Manhattan distance in synthetic Disguise database

#	Windows Combination	Disguise Types	Identification Accuracy (%)
1	5,21,25	Hair style	92.65
2	7,21,23,25,29	Glasses	91.67
3	13,19,25,31	Hat and cap	85.78
4	7,11,13,23	Beard and mustache	89.71
5	3,21,23,27,29	Eye and eyebrow	93.14
6	9,13,21,23	Aging	82.31
7	7,15	Multiple variation	69.71
8	5,9,13,15,23,25	All variations	84.72

3.3. The Effect of Reducing the Number of Gallery Images on the Identification Accuracy

The proposed face recognition algorithm, for the purpose of law enforcement and criminal justice applications where a limited number of gallery images is available, has been studied. The performance of the proposed algorithm is examined under conditions that the number of images reduces to the worst case scenario (In situations where there is only one image for each person); also the performance of the proposed face recognition algorithm has been compared with other available methods. Experiments have been performed on both databases described in the previous section. In the experiments with single gallery image, database contains a front view and Gray scale face image of each person and the rest of the images used as probe images. In the experiments with two or three images, gallery consists of Lightless images with minimal variations or disguise accessories and the rest of the images used as probe images. First the images must be converted to Gray scale, then the feature extraction is performed, finally matching operations will be done. Table 3 shows the identification accuracy of algorithms under the condition that the number of images reduces from three to one. As can be seen the proposed method has the best performance in the condition of three or two gallery images, and with single gallery image it has the comparable performance despite of its uncomplicated recognition process.

3.4. The Effect of Various Types of Disguise on the Identification Accuracy

Investigation of the performance of the proposed face recognition algorithm, when there are various types of disguise in images, is the subject of further study in this paper. All the experiments in this section are evaluated in order to determine the recognition accuracy for a variety of facial changes under the condition that there would only be one image of per person. In the experiments with single gallery image, database contains a front view and Gray scale face image of each person and the rest of the images used as probe images. By applying the SVD method to single gallery image, from each image, three renovated images are produced. This approach increases the efficiency of the proposed method.

Image databases are divided into some different categories according to disguise variations and proposed algorithm functionality is evaluated for each category. The table 4 shows the results of evaluation of the proposed method with variety of disguise on these databases and in comparison to the existing algorithms. According to these results, performance of the proposed algorithm is different for each types of the disguise. About AR database, those images with facial changes have the highest accuracy, because only a small area of the face is being affected and the remaining parts may help to have correct diagnosis. In the case of images with glasses covered areas of the face may result in decreased accuracy. In the database of dummy facial change, two groups of changes including eye / brow and hair style might produce the highest accuracy. In all categories of disguise the proposed method has brought about the competitive performance relative to the other algorithms.

4. Conclusion

In spite of numerous researches in the field of face recognition, the researchers still have many problems; which little number of gallery images for recognition task and disguise can be named. To face these challenges, in this paper an approach based on local phase quantization method and singular value decomposition (SVD) algorithm was proposed. The SVD method is being used to cope with the challenges of the images gallery limitation and also to increase the training samples. The results of this study indicated that the identification accuracy increased by using the combination of windows with different sizes in LPQ method. For this purpose, Brute Force algorithm applied to all possible combinations of windows, and a combination with the highest identification accuracy was selected as the optimal combination. This process was conducted separately for a variety of disguise and for each of them have different optimum combination was obtained. The ability of the proposed method was shown by performed experiments on AR and Synthetic Disguise databases. Because of unavailability of standard disguised database, we prepared a database particularly for disguise consist 3570 images from 85 persons in 42 different modes by using Faces software. The performance of the proposed algorithm was evaluated and compared with some available algorithms. Results showed the ability of the proposed algorithm in disguised face recognition compared to other algorithms. Also when we have only an image of the person, the proposed algorithm achieved the highest recognition accuracy.

Table 3

Identification accuracy of algorithms changing number of gallery image

Database	number of the gallery's pictures	Identification Accuracy (%)						
		*PCA**	*GF**	*LFA**	*IGF**	*LBP**	*2D-LPGNN**	*The proposed Method*
AR	Three	50.6	49.8	68.5	92.0	93.2	94.5	96.53
	Two	38.7	41.7	51.5	86.6	87.9	90.0	93.66
	One	28.4	36.4	40.6	77.1	78.7	81.2	78.52
Synthetic Disguise	Three	86.5	90.1	92.9	93.4	94.1	95.6	96.51
	Two	77.2	82.2	85.5	86.8	88.3	91.8	92.28
	One	69.6	79.0	80.2	80.6	81.4	83.2	84.72

Table 4

Identification accuracy of algorithms for disguise variations

Database	Disguise Types	Identification Accuracy (%)						
		*PCA**	*GF**	*LFA**	*IGF**	*LBP**	*2D-LPGNN**	*The proposed Method*
AR	Expressions with illumination	31.5	38.2	48.8	85.5	86.5	89.7	85.49
	Glasses with illumination	28.6	31.9	33.4	66.9	69.2	71.7	71.54
	Scarf with illumination	11.7	31.9	30.9	54.1	55.4	72.9	69.23
Synthetic Disguise	Hair style	85.7	94.1	94.8	94.8	94.8	94.9	92.65
	Glasses	70.9	55.1	84.4	84.6	84.9	85.2	91.67
	Hat/cap	82.6	90.9	91.7	92.8	93.1	94.7	85.78
	Beard and moustache	59.1	85.2	81.3	82.8	83.5	84.6	89.71
	Eye and eyebrow	87.4	78.6	96.3	96.7	96.8	97.1	93.14
	Aging	77.5	81.8	92.9	94.0	94.4	95.4	82.31
	Multiple	19.7	49.1	61.3	63.3	63.5	71.2	69.71

References

[1] Y. Zhai, J. Gan, J. Zeng, and Y. Xu, "Disguised face recognition via local phase quantization plus geometry coverage," in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 2013, pp. 2332-2336.

[2] R. Singh, M. Vatsa, and A. Noore, "Face recognition with disguise and single gallery images," Image and Vision Computing, vol. 27, pp. 245-257, 2009.

[3] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based face recognition using mixture-distance," in Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on, 1996, pp. 209-216.

[4] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of cognitive neuroscience, vol. 3, pp. 71-86, 1991.

[5] P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," Network: computation in neural systems, vol. 7, pp. 477-500, 1996.

[6] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," Neural Networks, IEEE Transactions on, vol. 8, pp. 98-113, 1997.

[7] B.-L. Zhang, H. Zhang, and S. S. Ge, "Face recognition by applying wavelet subband representation and kernel associative memory," Neural Networks, IEEE Transactions on, vol. 15, pp. 166-177, 2004.

[8] N. Ramanathan, R. Chellappa, and A. Roy-Chowdhury, "Facial similarity across age, disguise, illumination and pose," in Image Processing, 2004. ICIP'04. 2004 International Conference on, 2004, pp. 1999-2002.

[9] G. Golub, C. Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1983.

[10] Q-X. Zhang, L. Zhang, and D. Zhang, " Face recognition using FLDA with single training image per person, " Applied Mathematics and Computation 205 (2008) 726–734

[11] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in Image and signal processing, ed: Springer, 2008, pp. 236-243.

[12] C. Liu and H. Wechsler, "Independent component analysis of Gabor features for face recognition," Neural Networks, IEEE Transactions on, vol. 14, pp. 919-928, 2003.

[13] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, pp. 2037-2041, 2006.

[14] Available:http://www.iqbiometrix.com/products_faces_40.html

[15] A. Martınez and R. Benavente, "The AR face database," Computer Vision Center, Universitat Autónoma de Barcelona, CVC Technical Report, vol. 24, 1998.