



Efficient k-class approach for face recognition [☆]



Amin Alqudah ^{*}, Hussein Al-Zoubi

Computer Engineering Department, Hijjawi Faculty for Engineering Technology, Yarmouk University, Irbid 21163, Jordan

ARTICLE INFO

Article history:

Received 6 July 2014

Received in revised form 29 March 2015

Accepted 30 March 2015

Available online 16 April 2015

Keywords:

Face detection

Face recognition

Recognition rate

Self-organizing-maps

ABSTRACT

In this research work, a new k-class approach for efficient and accurate face recognition called kCAFRe is established. The kCAFRe system has three stages: preprocessing, training, and testing. In the training phase, four reference pictures (classes) are constructed for each person. During testing, the correlation coefficient (CC) is calculated between the picture under test and each of the reference pictures. For each person, one accredited class is chosen. Subsequently, the accredited class for the person with the highest CC is selected. Results were given for four image databases. A recognition ratio of 97% has been obtained for the Libor Spacek's set.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Face recognition is an important branch of biometrics, which does not need the cooperation of the test subjects to work properly. It is important for police and military purposes involving national and international security. For example, in pervasive computing, face recognition is used to support human memory like smart eyeglasses and wearable face recognition systems [1–3].

Although face recognition has greatly advanced in the last few years, it still faces many challenges, such as dealing with faces in low-quality pictures or with inadequate lighting. Ideally, face recognition systems should address the following issues: (i) to deal with non-controlled environments, (ii) to consider databases containing multiracial faces, and (iii) to take into account aging effects [4]. Surface deformations caused by facial expressions represent another challenge in 3-D face recognition [5]. Therefore, more research efforts are still needed to leverage the realm of face recognition to higher levels. Recently, researchers have shown more interest in face recognition (aka identification) than face verification (aka authentication) [4].

This work briefly reviews some of the most recent studies in the literature on this topic in order to help interested researchers find ongoing research on 2D and 3D face recognition and verification. Many studies have relied on Gabor-based methods for face biometrics including recognition and verification. Serrano et al. [4] divided Gabor-based methods for face biometrics into two categories: analytical and holistic. The graph-based [6–8] and non-graph-based [9–12] in addition to enhanced methods (using optimal Gabor parameters as well as Gabor and Adaboost) [13–16] are the most famous analytical methods. On the other hand, the holistic methods include downsampled Gabor [17,18] and [15], PCA/ LDA [19,20], kernel PCA/LDA [19] and [21], Gabor 2D methods [22–25], local binary patterns [26–28] and [8], and no downsampling methods [22,23].

[☆] Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. Zhanpeng Jin.

^{*} Corresponding author.

E-mail addresses: amin.alqudah@yu.edu.jo (A. Alqudah), halzoubi@yu.edu.jo (H. Al-Zoubi).

The authors of [29] represented each set of images as a subspace, applied a locally orthogonal method for subspace-to-subspace matching to improve accuracy along with online learning to reduce the computational cost. In the experiments of [29], a small video database was used and the achieved recognition rate was only 93.4%. In [26], the authors performed face identification and face verification on various databases including FERET, CAS-PEAL, CMU-PIE, Extended Yale B, and FRGC. Their proposed method for face recognition used local derivative pattern (LDP) based on high-order local derivative variations to encode directional pattern features. The reported recognition rates in [26] ranged from 78.9% to 97.9%. Nevertheless, the achievement of 97.9% recognition rate was only possible with the aid of manual alignment and cropping. The human face recognition proposed in [30] combined fast discrete curvelet transform (FDCT) and two-dimensional principal component analysis (2DPCA). The authors used FDCT as a feature extractor, while 2DPCA was applied to reduce the computational complexity. Complex wavelet moments (CWMs) can describe local and global characteristics of images, and therefore have been used by Singh and Sahan [31] as image descriptors for real-time face recognition. In their work, the authors used very large image sizes, performed experiments in part on a small database, and manually aligned and cropped images. The face recognition framework proposed in [32] integrates a local feature-based approach and a holistic approach.

The authors of [33] proposed to use a margin-based between-class scatter and a regularization process based on the support vector machine (SVM) margins for feature extraction. The recognition rates of [32,33] were relatively not very high. Furthermore, manual alignment and cropping have been employed. In [34], particle swarm optimization (PSO) was proposed to simultaneously optimize the parameters of support vector machine (SVM) for face recognition. The authors of [35] proposed a locality Weighted Sparse Representation based Classification (WSRC) method for face recognition, where they considered the Sparse Representation based Classification (SRC) as a generalization of Nearest Neighbor (NN) and Nearest Feature Subspace (NFS). The authors of [36] enhanced face recognition by using image pre-processing techniques like resolution conversion, histogram equalization, and edge detection and proposed a method called Accelerated Binary Particle Swarm Optimization (ABPSO) for feature selection. As another approach, super-resolution (SR) methods can be used to reconstruct a high resolution image from single or multiple low-resolution images. In [37], a method based on dictionaries in high frequency wavelet subbands was proposed using super resolution for enhancing face recognition. The authors in each of the papers [34–37] relied on manual alignment and cropping in their work in addition to using very large image sizes.

In [38], the authors used pseudo Zernike moments to extract features in their face recognition system. In [39], chain codes of face contours have been used for feature extraction. Lastly, the authors of [40] proposed combining Chi square and entropy for feature extraction and Chi square and entropy for classification. The authors in [38–40] relied on a small sample size (by using only part of the Libor Spacek's database) and the reported recognition ratios were not very high.

This paper is organized as follows: Section 2 introduces the proposed face recognition system, which consists of preprocessing, training, and testing. Next, in Section 3, the experimental results are provided and discussed. Finally, Section 4 concludes this paper.

2. The proposed face recognition system

The proposed face recognition system is called **k-Class Approach Face Recognition (kCAFRe)**, which is able of tackling k-classes. It consists of three major components (as illustrated by the block diagram of Fig. 1): the preprocessor, the trainer, and the tester. The details of each of these components will appear later in the coming subsections.

The first step in kCAFRe is preprocessing. In this step, the faces of the images (if exist) are detected and cropped automatically. The output of this face detection sub step is usually faces with different sizes. Therefore, the next sub step of the preprocessor is resizing the detected faces to a common size. Then, all images are converted to gray images and divided into two subsets; training subset, and testing subset.

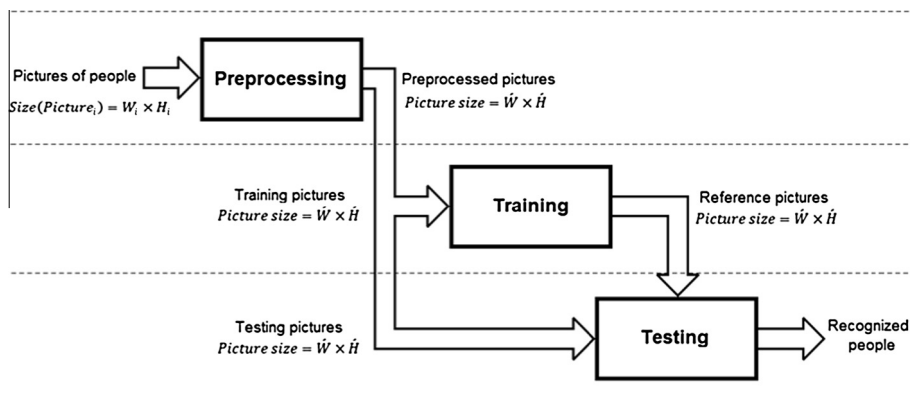


Fig. 1. The k-Class Approach Face Recognition (kCAFRe) system.

The second step in kCAFRe is training. In this step, the training subset of the images is used to train a Self Organizing Maps (SOM) classifier, where there is a classifying model of certain number of classes for each person. All people models are used to test images in the testing subset in the next step.

The final step of kCAFRe is testing. In this step, the testing subset of the images is tested for recognition using the classifying models created in the previous step. For each image under testing, there are different classes for each person in the database. The main feature that is used to decide the similarity to the testing image is the correlation. Therefore, the subclass that has the highest correlation value with the testing image is decided to be the recognition output.

2.1. The preprocessor

The preprocessor prepares input images for the next two steps: training and testing. For this, the preprocessor performs the following tasks using a MATLAB implementation:

- Face detection: The face of the person in the input picture can be located anywhere inside that picture. This step locates and extracts faces from input pictures using the Viola and Jones face detection algorithm (VJFDA) for face cropping [41,42].
- Image size reduction: The training and testing steps need the subregion of the picture that contains the face (face cropping). The output of the previous step is the part of the original image (a rectangle) that contains the face after resizing to produce a picture of a standard size.
- Color-to-gray picture conversion: For the sake of portability, kCAFRe works on gray-scale pictures. Therefore, regardless of the color system of the input picture, the preprocessor always produces gray-scale pictures.

The three main characteristics of a successful face recognition system are high recognition rates, high computational performance, and simplicity (which leads to an efficient implementation). kCAFRe has these characteristics as will be shown shortly. The image size reduction done by the preprocessor produces a fast face recognition system, because the trainer and the tester will operate on a lower number of pixels and, therefore, perform fewer processing operations. The size of an input picture i in pixels is (please refer to Table 1, that lists the symbols used in this study and their meanings):

$$\text{Size (Picture}_i) = W_i \times H_i. \quad (1)$$

The output of the preprocessor is a picture of standard size given by

$$\text{Picturesize} = \dot{W} \times \dot{H}, \quad (2)$$

where $\dot{W} \ll W_i$ and $\dot{H} \ll H_i$. The reduction factor between the input picture i and the standardized image is

$$\text{Reduction factor} = \frac{W_i \times H_i}{\dot{W} \times \dot{H}}. \quad (3)$$

Assuming that $\dot{W} = \dot{H} = \sqrt{A}$, Eq. (3) becomes

$$\text{Reduction factor} = \frac{W_i \times H_i}{A}. \quad (4)$$

Table 1

List of symbols used in this study and their meanings.

| Symbol | Meaning |
|-----------|--|
| W_i | Original width of input picture of Person $_i$ in pixels |
| H_i | Original height of input picture of Person $_i$ in pixels |
| \dot{W} | Width of preprocessed picture in pixels |
| \dot{H} | Height of preprocessed picture in pixels |
| A | Area of a square preprocessed picture |
| N | Total number of people in the database |
| P_i | Number of available pictures (models) for Person $_i$ |
| P | Number of pictures of each person under the assumption that all persons have the same number of images |
| f | Fraction of pictures used in training |
| Tr | Total number of pictures in the training subset |
| Ts | Total number of pictures in the testing subset |
| R | Number of reference images for each person |
| C | Total number of classes in the system |
| ϕ_k | Correlation coefficient (CC) between the test picture and the reference picture k |

2.2. Self-Organizing Maps (SOM)

In the jargon of neural networks, there are two techniques to train a network. The first one is the **supervised training** where each input pattern has a target and the network learns (changes its weights) to represent the relation between the input and the output patterns. The second training technique is the **unsupervised training**, where the network learns (changes its weights) to classify the input patterns without prior information [43].

In this paper, our interest is in the **unsupervised neural networks** and especially in those which are **based on competitive learning, where neurons compete to activate so that only one neuron (winning neuron) is activated at any time** [43]. By this activation process, neurons organize themselves. Therefore, this kind of network is named “Self Organizing Map” (SOM). Fig. 2 shows the two dimensional SOM structure [44].

The main goal of a SOM is to adaptively perform a **nonlinear transformation or mapping of data onto a one or two-dimensional map**. In order to do so, the network has to train (change its weights) based on the following steps [43]:

- (1) **Initialization:** In this step, the connection weight vectors $\{\mathbf{w}_j, j \in [1, M], M \text{ is the number of neurons}\}$ of all neurons are initialized randomly.
- (2) **Sampling:** In this step, an input pattern \mathbf{x} is applied.
- (3) **Matching:** In this step, the winning neuron $i(\mathbf{x})$, is decided. The winning neuron is the one that has the weight vector closest to the applied input pattern. For this purpose, Euclidean distance is calculated.
- (4) **Updating:** In this step, the weight vectors of the winning neuron and its neighbor neurons are updated based on the following equation:

$$\mathbf{w}_j = \mathbf{w}_j + \eta h_{j,i(\mathbf{x})} (\mathbf{x} - \mathbf{w}_j) \quad (5)$$

where η is the learning rate; h is the Gaussian neighborhood function, and is defined as

$$h_{j,i(\mathbf{x})} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right) \quad (6)$$

where $d_{j,i}$ is the lateral distance between the winning neuron and neuron j , and σ is the width of the Gaussian neighborhood function.

- (5) **Continuation:** Return to step 2 and repeat until weights stop changing.

2.3. The trainer

Assuming that N represents the total number of people in the database and P_i represents the number of available pictures for Person _{i} . If all persons have the same number of images i.e.

$$P_1 = P_2 = \dots = P_N = P. \quad (7)$$

Then, the total number of pictures available in the database can be calculated by

$$\text{Database size} = N \times P. \quad (8)$$

The set of $N \times P$ pictures is divided into two subsets: the training subset and the testing subset. The size of training subset is

$$Tr = f \times N \times P, \quad (9)$$

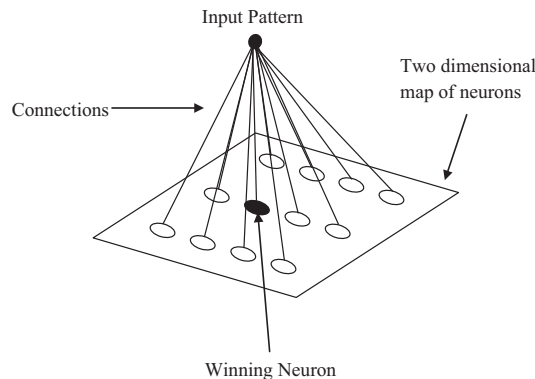


Fig. 2. Two dimensional SOM structure [44].

where f ranges between 0% and 100%. The remaining pictures are used for testing, where the size of the testing subset is

$$T_s = (1 - f) \times N \times P. \quad (10)$$

In this work, $f = 65\%$; which means, **65% of the total images was used for training, and 35% was used for testing**. The proposed kCAFRe face recognition system adopts a k -class approach. Therefore, the trainer builds for each person a number of reference pictures (classes) R , where $R \leq Tr$. The total number of classes in the system is

$$C = N \times R. \quad (11)$$

The construction of each of the reference pictures is discussed next. The Tr pictures for each person are classified into R classes using SOM. In this work, the optimal value of R was found to be 4 empirically; for this number of classes, the recognition rate was found to be the highest. After classification, each class can have any number of pictures with the restriction that each class should have at least one picture. The reference picture of class k is the average picture of all the pictures that belong to class k . That is, the value of Pixel_{ij} where i and $j \in [1, A]$ of reference k is

$$\text{Pixel}_{ij}(\text{Reference Picture}_k) = \frac{\sum_{l=1}^m \text{Pixel}_{ij}(\text{Picture}_l)}{m} \quad (12)$$

assuming that class k contains m pictures. If $\text{Pixel}_{ij}(\text{Reference Picture}_k)$ is a non-integer value, then rounding toward nearest integer is applied.

The reference pictures are constructed offline and kept in a separate database to be used in the testing phase.

2.4. The tester

Given a test picture, the tester goes over each reference picture and calculates the CC feature ϕ between the test picture and the reference picture using

$$\phi = \frac{\frac{1}{A} \sum_{i=1}^A (I_R(i) - m_{I_R})(I_T(i) - m_{I_T})}{\sigma_{I_R} \sigma_{I_T}}, \quad (13)$$

where $I_R(i)$ and $I_T(i)$ are the intensity values for pixel i in the lexicographically ordered vectors representing the reference and the test images, respectively. A is the number of pixels in the image, m_{I_R} , m_{I_T} , σ_{I_R} and σ_{I_T} are the mean and standard deviation of the reference and test images, respectively.

For each person, only one class is chosen (the class that has the highest correlation value). The chosen class is called *accredited class*. Finally, the person that has the highest correlation value of its accredited class is decided to be the recognition output. The operation of the tester is explained by [Algorithm 1](#).

Algorithm 1. Testing a picture

Input: Test picture,
Database of $N \times R$ reference pictures.
Output: Recognition decision (index of a person).
 Initialize Decision = 1;
 $\phi_{\text{Highest}} = 0$;
 for Person _{i} , $1 \leq i \leq N$, Do:
 $\phi_{\text{Accreditted}}(\text{Person}_i) = 0$;
 for Class _{k} , $1 \leq k \leq R$, Do:
 Calculate ϕ_k ;
 if $\phi_k > \phi_{\text{Accreditted}}(\text{Person}_i)$, then
 $\phi_{\text{Accreditted}}(\text{Person}_i) = \phi_k$;
 end if
 end for
 if $\phi_{\text{Accreditted}}(\text{Person}_i) > \phi_{\text{Highest}}$, then
 $\phi_{\text{Highest}} = \phi_{\text{Accreditted}}(\text{Person}_i)$;
 Decision = i ;
 end if
 end for
 Return Decision.

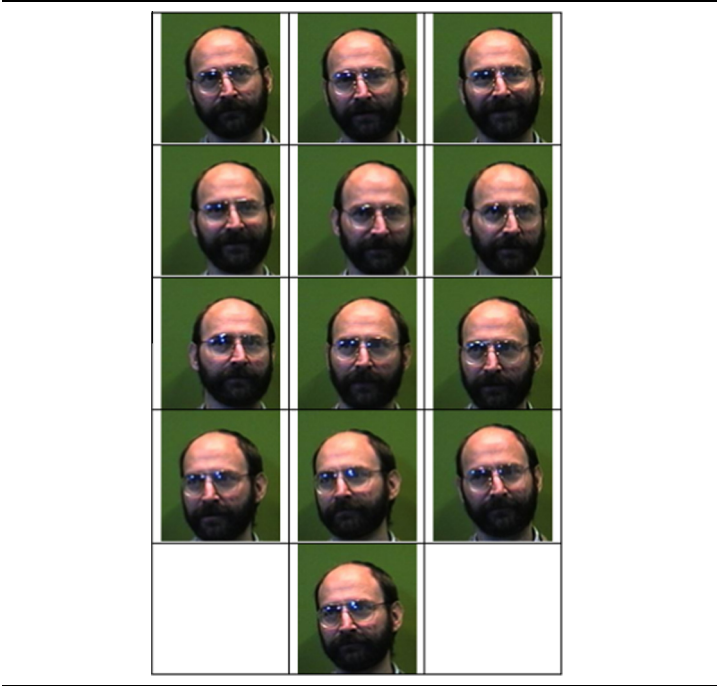
3. Evaluation results

The specifications of the computer used in the experiments are as follows: PC with Intel Core 2 Duo CPU at 2.1 GHz, 4 GB RAM, and 64-bit Microsoft Windows 7 OS. The Libor Spacek’s Facial Images database [45] was used to evaluate kCAFRe. Libor Spacek’s set contains four databases with pictures of 395 female and male distinct multiracial people of different ages: faces94, faces95, faces96, and grimace, where the difficulty increases in order in these databases in terms of background and scale, and variation of expressions. For each person in these databases there exist 20 pictures of different poses and with various face expressions for a total of 7900 pictures. Some people in the database wear eyeglasses and some have beards. It should also be pointed out that these 7900 pictures have different lighting conditions. In this work, 7240 pictures have been used for 362 individuals from the databases. The purpose in this research is to propose a fully automated face recognition system. Therefore, any manual cropping has been avoided. The reason for not using all the available pictures is VJFDA for face detection and cropping. The total number of pictures that VJFDA was able to detect and crop was 7240 pictures for 362 individuals. Table 2 lists the values of the various parameters used by the experiments that evaluate kCAFRe results.

Table 2
Values of the various parameters used in this study.

| Parameter | Value |
|-------------|----------|
| W_i | Variable |
| H_i | Variable |
| \acute{W} | 25 |
| \acute{H} | 25 |
| A | 625 |
| N | 362 |
| P_i | Variable |
| P | 20 |
| f | 0.65 |
| Tr | 4706 |
| Ts | 2534 |
| R | 4 |
| C | 1448 |

Table 3
The 13 training pictures for an example person. These pictures are in their original format before cropping or preprocessing, with an equal size 180×200 pixels. The person shown in the pictures has a beard and wears eyeglasses.



As shown in Table 2, $f = 0.65$, therefore, for each person, 13 pictures will be used for training and the remaining 7 pictures are for testing. The pictures used in the training phase are picked at random. The 13 training pictures for an example person are shown in Table 3. These pictures are in their original format (before cropping or preprocessing), with an equal size 180×200 pixels. In Table 4, the pictures of Table 3 are shown in corresponding order after face detection using VJFDA and after automatic cropping. It is natural that these pictures differ in size because the location and size of the face varies from picture to picture. It has been noticed that the average face size captured by VJFDA was around 100×100 pixels.

Table 4
The 13 training pictures for an example person of Table 3 after face detection using VJFDA and after cropping. Note that these pictures differ in size.






























| | | |
|---|---|---|
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| |  | |

Table 5
The 13 training pictures for an example person of Table 4 after resizing to 25×25 pixels.

| | | |
|---|---|---|
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| |  | |

| | | | |
|---|---|---|---|
|  |  |  |  |
|---|---|---|---|

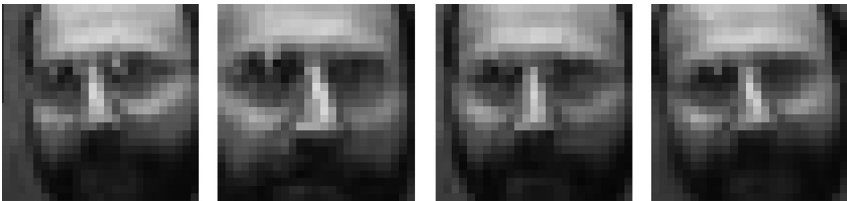


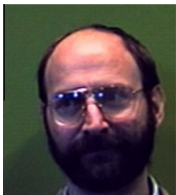
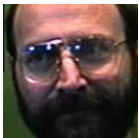


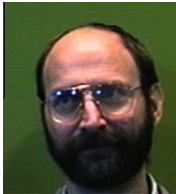
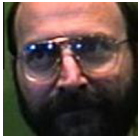


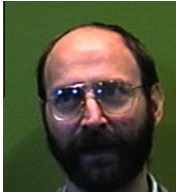
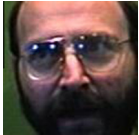


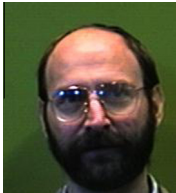



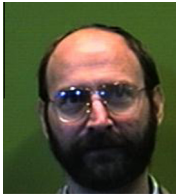



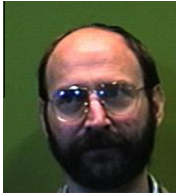
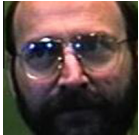


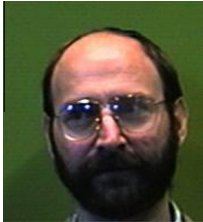
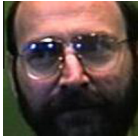


Fig. 3. The four SOM class pictures calculated for the example person.

Table 5 shows the pictures of Table 4 after resize to 25×25 pixels in the corresponding order. The four SOM class pictures calculated for this example person are shown in Fig. 3.

In Table 6, the seven test pictures for the example person are shown. The first column depicts the picture's name as shown originally in the database. The second column shows the original picture before preprocessing. The third column displays

Table 6

The seven test pictures for the example person.

| Name | Before crop 180×200 | After crop | After resize 25×25 | Accredited class | Class # | Correlation |
|---------------|---|---|---|---|---------|-------------|
| anonym2.3.jpg |  |  |  |  | 4 | 0.9563 |
| anonym2.4.jpg |  |  |  |  | 4 | 0.9648 |
| anonym2.5.jpg |  |  |  |  | 4 | 0.9520 |
| anonym2.6.jpg |  |  |  |  | 4 | 0.9856 |
| anonym2.7.jpg |  |  |  |  | 4 | 0.9886 |
| anonym2.8.jpg |  |  |  |  | 4 | 0.9831 |
| anonym2.9.jpg |  |  |  |  | 4 | 0.9743 |

pictures after face detection using VJFDA and after cropping. The fourth column shows the picture after resize to 25×25 pixels. The fifth column shows the accredited class when testing that specific picture. The sixth column shows the number of the accredited class, and the last column shows the CC value between the picture under test and the accredited class. It can be noticed that for this example person the accredited class for all these seven test pictures happened to be class 4.

The total time required to accomplish each of the major phases of this work is shown in Table 7. These times represent the elapsed times to process all the pictures in the database.

As shown in Table 2, the total number of pictures used in the testing phase was 2534. kCAFRe was successful in correctly recognizing 2457 faces, which means a recognition rate of 97%. The time to test all the images was 28,733.72 s. Which means the system would need about 11.3 s to recognize any image among 2534 images. This time is quite comparable to other systems as will be shown in Section 3.

The achieved identification rates of kCAFRe are shown in Fig. 4 for Ranks 1–10. The achieved identification rates lie in the range 97–99.5%. Fig. 5 shows the achieved verification rates of kCAFRe using false acceptance rates ranging from 0.001 to 1. The achieved verification rates range between 83% and 100%.

3.1. Testing using different similarity measures

In this part, kCAFRe is further investigated using another two similarity measures. The first one is Structural SIMilarity (SSIM) [46,47] and the second one is based on Fast Fourier Transforms. Below, a brief description of each of the two measures is provided with the achieved recognition rates when applying kCAFRe based on them.

3.2. SSIM metric

SSIM index is a metric used to measure the similarity between two images. It is a modified metric of the universal image quality index [48,49]. SSIM takes two images and compares the local patterns of pixel intensities. These patterns have to be

Table 7
The total time required to accomplish each of the major phases.

| Procedure | Face detection | Picture resize | Training | Testing |
|-------------|----------------|----------------|----------|----------|
| Time (sec.) | 22,285.61 | 1580.33 | 2172.25 | 28733.72 |

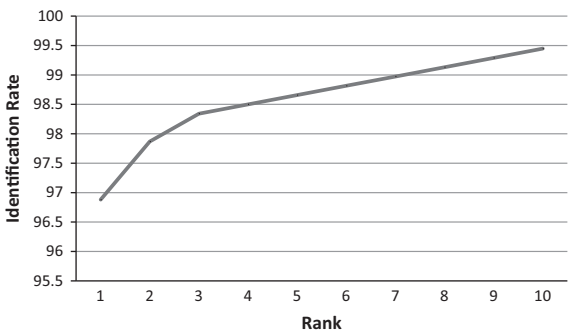


Fig. 4. The achieved identification rates.

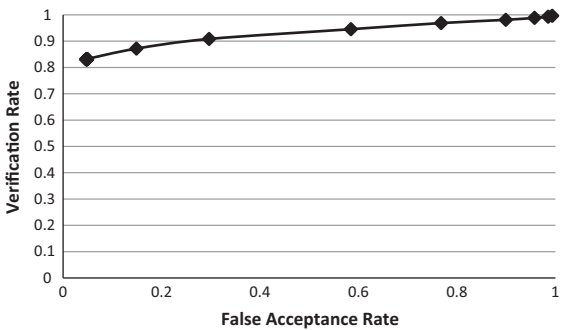


Fig. 5. The achieved ROC values.

normalized for luminance and contrast. The following equations show how this index is calculated to measure the similarity between two images (x and y) [47].

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (14)$$

where

$$C_1 = (K_1L)^2, \quad (15)$$

$$C_2 = (K_2L)^2, \quad (16)$$

and K_1 is the a small constant (0.01); K_2 the a small constant (0.03); L : dynamic range of pixel values (255 for grayscale images); μ_x the mean of image x ; μ_y the mean of image y ; σ_x^2 the variance of image x ; σ_y^2 the variance of image y ; σ_{xy} is the covariance between images x and y .

In this paper, the SSIM index was used to measure the similarity between the images under investigation and all people classes in the database. SSIM values ranges between '0' and '1'. SSIM value of '0' means no similarity between the two images, while SSIM value of "1" means images x and y are similar (identical). As mentioned before, the total number of pictures used in the testing phase was 2534. This similarity metric was successful in correctly recognizing 2464 faces, which means a recognition rate of 97.23%.

3.3. FFT-based metric

The second metric is based on FFT. If we have two images x and y , this metric is based on finding the distance (D) between X and Y , where X and Y are the 2-dimensional FFT representation of x and y , respectively as shown below [50].

$$X(u, v) = \sum_{r=0}^{\hat{H}-1} \sum_{c=0}^{\hat{W}-1} x(r, c) e^{-j2\pi \left(\frac{ur}{\hat{H}} + \frac{vc}{\hat{W}} \right)}, \quad (17)$$

$$Y(u, v) = \sum_{r=0}^{\hat{H}-1} \sum_{c=0}^{\hat{W}-1} y(r, c) e^{-j2\pi \left(\frac{ur}{\hat{H}} + \frac{vc}{\hat{W}} \right)}, \quad (18)$$

where $u = 0, 1, 2, \dots, \hat{H} - 1$, $v = 0, 1, 2, \dots, \hat{W} - 1$, \hat{H} is the number of rows in the image, and \hat{W} is the number of columns in the image.

$$D = \left(\sum_{u=0}^{\hat{H}-1} \sum_{v=0}^{\hat{W}-1} (X(u, v) - Y(u, v)) \cdot (X(u, v) - Y(u, v))^* \right)^{1/2} \quad (19)$$

In this paper, the distance D was calculated to recognize the images under investigation. If we have two images x and y and their FFT representations is X and Y , respectively. If the two images are very similar to each other, then the distance D between X and Y will be close to "0". On the other hand, if the two images are very different, then the distance D between X and Y will be far from "0". Therefore, the image under investigation was recognized based on calculating the distance between its FFT representation and the FFT representation of all the other classes. The person who has the smallest distance of its accredited class is decided to be the recognition output. This similarity metric was successful in correctly recognizing 2469 faces, which means a recognition rate of 97.43%.

In summary, three similarity measures have been tried in this paper. The results shown earlier are based on the CC similarity measure because it is the simplest. The other two measures got almost the same recognition rate ~97%.

3.4. Man vs. machine

In this section, an extra experiment was carried out. All the misrecognized pictures were given to 213 people. Each participated individual was asked to decide between two choices which one is closer to the misrecognized picture. The first choice was what the proposed model recognized the image as, and the second choice was the picture of the person that should have been recognized as. Participants were not told which choice is the true choice. The participants' recognition rate for the misrecognized images was 81.1%.

It is worth mentioning that in this experiment, the participants (humans) were put in circumstances that make the recognition task easy. In the experiment, the participants were asked to decide between only two choices instead of finding a match in a large database. In spite of this, a large portion of the participated individuals chose the wrong answer. It is important to point out that only 3 out of 213 persons were able to correctly recognize all the images.

3.5. Comparison with the state-of-the-art methods

In this subsection, kCAFRé is compared with the state-of-the-art methods. It is in fact a tough matter to fairly compare between two or more face recognition methods. Let us illustrate this by raising the following situation: let us assume two recognition methods, Method 1 achieved a recognition rate of 99% on a database of size 100 pictures, while Method 2 achieved a recognition rate of 96% on a database of size 1000 pictures. Which method is better? In light of this, the authors think a “good” face recognition method should possess the following features:

1. Be fully automatic: the good face recognition method should not need any human intervention like for example manual cropping.
2. Work on a large database consisting of large number of individuals: the larger the number of pictures used to test a face recognition method, the better the method.
3. Achieve high identification rates.
4. Achieve high verification rates.
5. Use off-the-shelf components.
6. Be fast: in real-life applications, a fast face recognition method is highly desirable. Preprocessing, training, and testing should take minimal delay.

Table 8

Comparison between the state-of-the-art face recognition methods and the proposed method.

| Method | Databases used | T_r | T_s | A pixels | Time to test one face (sec.) | Recg. rate | Manual alignment/cropping |
|------------------------------|---|---------------------------------|---------------------------------|------------------|------------------------------|------------|---------------------------|
| Pang et al. [38] | Libor Spacek's database (Face94 + Face95) | 100 subjects \times 10 Images | 100 subjects \times 10 Images | 180 \times 200 | – | 97.51 | |
| | | 72 subjects \times 10 Images | 72 subjects \times 10 Images | 180 \times 200 | | 91.73 | |
| Boodoo-Jahangeer et al. [39] | Libor Spacek's database (Face94) | 20 subjects \times 20 Images | | 180 \times 200 | – | 95 | |
| Kakarwal and Ratnadeep [40] | Libor Spacek's database (Face94) | 10 subjects \times 15 Images | 10 subjects \times 5 Images | 180 \times 200 | – | 94 | |
| Zhang et al. [26] | FERET | 3541 | | 88 \times 88 | 0.054 | ~90% | |
| | CAS-PEAL | 3109 | | 88 \times 88 | – | ~83% | |
| | CMU-PIE | 68 subjects \times 1 Image | 68 subjects \times 146 Images | 88 \times 88 | – | 78.9% | |
| Singh and Sahan [31] | Extended Yale B | 38 subjects \times 9 Images | | 84 \times 96 | – | 97.9% | Yes |
| | JAFFE | 10 subjects \times 21 Images | | 256 \times 256 | 4.4 | 99.9% | |
| | FERET | 14,051 | | 256 \times 384 | 36.45 | 91% | Yes |
| | UMIST | 20 subjects \times 11 Images | 344 | 220 \times 344 | 10.6 | 97.6% | |
| Geng and Jiang [32] | ORL | 40 subjects \times 6 Images | 40 subjects \times 4 Images | 92 \times 112 | – | 96.2% | Yes |
| | AR | 75 subjects \times 7 Images | 75 subjects \times 7 Images | 60 \times 85 | – | 95.43% | |
| | GT | 50 subjects \times 8 Images | 50 subjects \times 7 Images | 60 \times 80 | – | 92.57% | |
| Kim et al. [33] | ORL | 40 subjects \times 5 Images | 40 subjects \times 5 Images | 50 \times 57 | – | 95.5% | Yes |
| | FERET | 1702 | | 56 \times 46 | – | <92% | Yes |
| | AR | 120 subjects \times 14 Images | | 56 \times 46 | – | <92% | Yes |
| Aneesh et al. [36] | CMU-PIE | 1840 | | 56 \times 46 | – | <92% | Yes |
| | Cambridge ORL | 40 subjects \times 4 Images | 40 subjects \times 6 Images | 92 \times 112 | – | 95.7% | Yes |
| AL-Hassan et al. [37] | Yale B | 38 subjects \times 22 Images | 38 subjects \times 36 Images | 168 \times 192 | – | 97.7% | Yes |
| | Extended Yale B | 38 subjects \times 15 Images | 38 subjects \times 49 Images | 168 \times 192 | – | 92.21% | Yes |
| kCAFRé | Libor Spacek's database | 362 subjects \times 13 Images | 362 subjects \times 7 Images | 25 \times 25 | 11.3 | 97% | No |

7. Use the minimal picture size: the larger the picture size, the better the chance to correctly recognize the person. However, the larger the picture size, the slower the method is because it operates on a greater number of pixels.

A comparison between the most relevant most recent state-of-the-art face recognition methods and kCAFRe is shown in Table 8. By investigating Table 8, it can be seen that none of these state-of-the-art methods provided verification results. Moreover, it can be noticed that many of state-of-the-art methods worked on a small-sized database, did not mention timing requirements, used large pictures sizes, or invoked manual face detection and cropping. The following shows detailed comparison between kCAFRe and those methods:

- In Zhang et al. [26] (FERET database): the time to test an image was 0.054 s. This time looks small. However, it was measured as the time for one-to-one image matching, which means if all the test images during the test were included in the testing phase, this time would increase exponentially. It is worth mentioning that the recognition rate of this algorithm is around 90% even though the image size was 88×88 pixels, which is much larger than the image size used in kCAFRe (25×25 pixels). Nevertheless, kCAFRe achieved higher recognition rate (97%).
- In Zhang et al. [26] (Extended Yale B database): few high resolution images were used to test the method. There are 38 subjects to be recognized, each has 9 images. Each image has a resolution of 84×96 pixels. The method achieved a high recognition rate (97.9%), but all the images that were used to test and train the system were manually aligned and cropped which leads to high recognition rates. Similar scenarios happened in Wei et al. [34] (FERET database), Aneesh et al. [36] (Yale B database) and Lu et al. [35] (Extended Yale B database), where the number of images was small and the images were manually aligned and cropped with high resolution. This explains the achieved high recognition rates.
- In Singh and Sahan [31] (JAFPE and UMIST databases): the algorithm achieved very high recognition rates (99.9% and 97.6%, respectively). But what helps the algorithm to get these values was the very high resolution images used in the training and testing phases. The images that were used have at least 100 times more resolution than the images used in this paper. This leads to a longer time to recognize any image. A comparison between these two algorithms with the proposed one would raise the following points:
 - The images in these two methods have at least 100 times more resolution than the images used in this paper.
 - The number of images used in training and testing in kCAFRe is $35 \times$ greater than the numbers used in these two methods. This would give the proposed method better generalization and better knowledge.
 - kCAFRe used more images than what the two methods used, where more time is needed to process a large database. Nevertheless, the timing of the proposed method is not far from theirs.

It is important to emphasize that the Libor Spacek's database is categorized into four categories: Face94, Face95, Face96 and Grimace. These categories are sorted from the easiest to recognize to the hardest, with Face94 being the easiest and Grimace the hardest category of faces to recognize.

It can be seen from Table 8 that the Libor Spacek's database has been used in [38–40]. However only part of the Libor Spacek's database has been in these papers (in fact the easiest categories of the Libor Spacek's database). In [38], only Face94 and Face95 have been tested and the reported recognition rates were 97.51% for Face94 and 91.73% for Face95. In [39], the simplest category of the Libor Spacek's database (that is Face94) has been used, where they scored a recognition ratio of 95%. Likewise, only the Face94 category has been experimented in [40], where a recognition ratio of 94% was achieved.

4. Conclusions and discussion

In this study, a novel k-class approach for efficient and accurate face recognition called kCAFRe is introduced. The kCAFRe system consists of three stages: preprocessing, training, and testing. In the training phase, four reference pictures (classes) are constructed for each person. During testing, the correlation coefficient is calculated between the picture under test and each of the reference pictures. For each person, one accredited class is chosen. Then, the accredited class for the person with the highest CC value is chosen. Results were given for the Libor Spacek's facial image database. A recognition ratio of 97% has been obtained for this set. kCAFRe has the following properties:

1. It is fully automatic: every step in the proposed method is automatically performed.
2. It was tested on four large databases consisting of 7240 pictures for 362 individuals.
3. It achieves high identification rates of 96.9% for Rank 1 and 99.5% for Rank 10.
4. It achieves high verification rates ranging between 83% and 100% using false acceptance rates ranging from 0.001 to 1, respectively.
5. It uses off-the-shelf components: VJFDA implementation in MATLAB for face detection and cropping.
6. It is fast. As shown in Table 7, it takes only 11.3 s to test and recognize a picture from a large database.
7. It requires small picture sizes of only 25×25 pixels.

References

- [1] Kurze M, Roselius A. Smart Glasses linking real live and social network's contacts by face recognition. In: AH'11 conference, March 12–14, 2011. Tokyo, Japan.
- [2] Wang H, Bao X, Choudhury R, Nelakuditi S. InSight: recognizing humans without face recognition. In: ACM HotMobile'13, February 26–27, 2013. Jekyll Island (Georgia, USA).
- [3] Utsumi Y, Kato Y, Kunze K, Iwamura M, Kise K. "Who are you? – a wearable face recognition system to support human memory. In: 4th augmented human international conference (AH'13), March 07–08 2013. Stuttgart, Germany. p. 150–3.
- [4] Serrano A, de Diego I, Conde C, Cabello E. Recent advances in face biometrics with Gabor wavelets: a review. *Pattern Recogn Lett* 2010;31:372–81.
- [5] Alyüz N, Gökberk B, Akarun Le. Regional registration for expression resistant 3-D face recognition. *IEEE Trans Inform Forensics Secur* 2010;5(3):425–40.
- [6] Lades M, Vorbrüggen JC, Buhmann J, Lange J, von der Malsburg C, Würtz RP, et al. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans Comput* 1993;42(3):300–11.
- [7] Qing L, Shan S, Chen X, Gao W. Face recognition under varying lighting based on the probabilistic model of Gabor phase. In: Proc. IAPR ICPR, vol. 3; 2006. p. 1139–42.
- [8] Zhang B, Shan S, Chen X, Gao W. Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition. *IEEE Trans Image Process* 2007;16(1):57–68.
- [9] Qin J, He Z-S. A SVM face recognition method based on Gabor-featured key points. In: Proc int conf machine learning and cybernetics, vol. 8; 2005. p. 5144–9.
- [10] Zou J, Ji Q, Nagy G. A comparative study of local matching approach for face recognition. *IEEE Trans Image Process* 2007;16(10):2617–28.
- [11] Jahanbin S, Bovik AC, Choi H. Automated facial feature detection from portrait and range images. In: Proc IEEE southwest symp image anal interpretation; 2008. p. 25–8.
- [12] Ilonen J, Kamarainen J-K, Paalanen P, Hamouz M, Kittler J, Kälviäinen H. Image feature localization by multiple hypothesis testing of Gabor features. *IEEE Trans Image Process* 2008;17(3):311–25.
- [13] Moreno P, Bernardino A, Santos-Victor J. Gabor parameter selection for local feature detection. In: Proc Iberian conf pattern recognition image anal. LNCS, vol. 3522. Springer-Verlag; 2005. p. 11–9.
- [14] Yang P, Shan S, Gao W, Li SZ, Zhang D. Face recognition using ada-boosted Gabor features. In: Proc IEEE int conf automatic face and gesture recognition; 2004. p. 356–61.
- [15] Shan S, Yang P, Chen X, Gao W. AdaBoost Gabor fisher classifier for face recognition. In: Proc int workshop anal modelling of faces and gestures. LNCS, vol. 3723. Springer-Verlag; 2005. p. 279–92.
- [16] Gao Y, Wang Y, Feng X, Zhou X. Boosting Gabor feature classifier for face recognition using random subspace. In: Proc IEEE ICASSP, vol. 2; 2006. p. II-II.
- [17] Liu C, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans Image Process* 2002;11(4):467–76.
- [18] Su Y, Shan S, Chen X, Gao W. Hierarchical ensemble of Gabor fisher classifier for face recognition. In: Proc IEEE int conf automatic face and gesture recognition; 2006. p. 91–6.
- [19] Liu C. Gabor-based kernel PCA with fractional power polynomial models for face recognition. *IEEE Trans Pattern Anal Machine Intell* 2004;26(5):572–81.
- [20] Xie X, Lam K-M. Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image. *IEEE Trans Image Process* 2006;15(9):2481–92.
- [21] Shen L, Bai L. Gabor wavelets and kernel direct discriminant analysis for face recognition. In: Proc IAPR ICPR, vol. 1; 2004. p. 284–7.
- [22] Serrano Á, Martín de Diego I, Conde C, Cabello E, Shen L, Bai L. Fusion of support vector classifiers for parallel Gabor methods applied to face verification. In: Proc MCS conf. LNCS, vol. 4472. Springer-Verlag; 2007. p. 141–50.
- [23] Serrano Á, Martín de Diego I, Conde C, Cabello E, Shen L, Bai L. Influence of wavelet frequency and orientation in an SVM-based parallel Gabor PCA face verification system. In: Proc IDEAL conf. LNCS, vol. 4881. Springer-Verlag; 2007. p. 219–28.
- [24] Jin Y, Ruan Q-Q. Gabor-based improved locality preserving projections for face recognition. In: Proc of IEEE ICIP, vol. 1; 2007. p. 153–6.
- [25] Wang L, Li Y, Wang C, Zhang H. Face recognition using Gabor face-based 2DPCA and (2D) 2 PCA classification with ensemble and multichannel model. In: Proc IEEE symp comput intell in security and defense appl; 2007. p. 1–6.
- [26] Zhang B, Gao Y, Zhao S, Liu J. Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE Trans Image Process* 2010;19(2):533–44.
- [27] Zhang W, Shan S, Gao W, Chen X, Zhang H. Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: Proc IEEE ICCV, vol. 1; 2005. p. 786–91.
- [28] He L, Hu D, Jiang C. Gabor binary codes for face recognition. In: Proc. Mexican int conf artificial intell; 2006. p. 53–60.
- [29] Kim T-K, Kittler J, Cipolla R. On-line learning of mutually orthogonal subspaces for face recognition by image sets. *IEEE Trans Image Process* 2010;19(4):1067–74.
- [30] Zhao J, Zhou Z, Cao F. Human face recognition based on ensemble of polyharmonic extreme learning machine. *Neural Comput Appl* 2013.
- [31] Singh C, Sahana A. Face recognition using complex wavelet moments. *Optics Laser Technol* 2013;47:256–67.
- [32] Geng C, Jiang X. Fully automatic face recognition framework based on local and global features. *Mach Vision Appl* 2013;24:537–49.
- [33] Kim S-K, Park Y, Toh K-A, Lee S. SVM-based feature extraction for face recognition. *Pattern Recogn* 2010;43:2871–81.
- [34] Wei J, Jian-qi Z, Xiang Z. Face recognition method based on support vector machine and particle swarm optimization. *Expert Syst Appl* 2011;38:4390–3.
- [35] Lu C-Y, Min H, Gui J, Zhu L, Lei Y-K. Face recognition via weighted sparse representation. *J Vis Commun Image Represent* 2013;24:111–6.
- [36] Aneesh MU, Masand A, Manikantan K. Optimal feature selection based on image pre-processing using accelerated binary particle swarm optimization for enhanced face recognition. In: International conference on communication technology and system design, procedia engineering, vol. 30; 2012. p. 750–8.
- [37] AL-Hassan N, Jassim S, Sellahewa H. Enhancing face recognition at a distance using super resolution. *MM&Sec'12*, coventry, September 6–7, 2012. United Kingdom. p. 123–32.
- [38] Pang Ying-Han, Teoh Andrew BJ, Ngo David CL. A discriminant pseudo Zernike moments in face recognition. *J Res Pract Inform Technol* 2006;38(2):197.
- [39] Boodoo-Jahangeer Nazmeen B, Baichoo Sunilduth. Face recognition using chain codes. *J Signal Inform Process* 2013;4:154.
- [40] Kakarwal Sangeeta N, Deshmukh Ratnadeep R. Hybrid feature extraction technique for face recognition. *Int J Adv Comput Sci Appl* 2012;3(2).
- [41] Viola Paul, Michael J. Jones. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, vol. 1; 2001. p. 511–8.
- [42] <http://www.mathworks.com/help/vision/ref/vision.cascadeobjectdetector-class.html>.
- [43] Haykin Simon. Neural networks: a comprehensive foundation. 2nd ed. Prentice Hall; 1998. ISBN 0132733501.
- [44] Kohonen Teuvo. The self-organizing map. *Proc IEEE* 1990;78(9):1464–80.
- [45] <http://cswww.essex.ac.uk/mv/allfaces/>.
- [46] <http://www.cns.nyu.edu/lcv/ssim/>.
- [47] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 2004;13(4):600–12.

- [48] Wang Z. Rate scalable Foveated image and video communications [Ph.D. dissertation]. Dept Elect Comput Eng. Austin (TX): Univ Texas at Austin; December 2001.
- [49] Wang Z, Bovik AC. A universal image quality index. *IEEE Signal Process Lett* 2002;9:81–4.
- [50] Lim Jae S. Two-dimensional signal and image processing. Englewood Cliffs, NJ: Prentice Hall; 1990. 710 p. 1 (1990).

Amin Alqudah received his MSE and Ph.D. in Electrical Engineering from the University of Colorado at Denver-USA in 2005, and Colorado State University-USA in 2009, respectively. Since 2009, he has been working with the Department of Computer Engineering, Yarmouk University-Jordan. He is currently an Assistant Professor. His research interests include artificial intelligence, pattern recognition and image processing.

Hussein Al-Zoubi received his MSE and Ph.D. in Computer Engineering from the University of Alabama in Huntsville, USA in 2004 and 2007, respectively. Since 2007, he has been working with the Department of Computer Engineering, Yarmouk University, Jordan. He is currently an Associate Professor and Department's chair. His research interests include machine vision, pattern recognition, image processing, and computer networks.