Gender Estimation using Single Exemplar of Frontal Facial Image per Person

Rohollah Akbari[1] and Saeed Mozaffari[2]

[1]*Electrical and Computer Department, Islamic Azad University of Qazvin, Qazvin, Iran*
[2]*Electrical Department, Semnan University, Semnan, Iran*
r_akbari@yahoo.com, mozaffari@semnan.ac.ir

Abstract

So far, several methods have been proposed for gender estimation based on training processes, using several images per person. In this paper, we introduce a new gender estimation method based on single image per person. First, reference images, called general faces, are constructed by averaging male and female images in the dataset respectively. Then, important face features such as eyes and mouth are highlighted by Wavelet fusion of Sobel Detector, and Laplacian of Gaussian images. The highlighted image is segmented into seven partitions. In the obtained image, those segments like eyes and eyebrows with more variations (containing more information) play more significant role in the gender estimation process. Weight of each segment is proportional to its entropy. Finally, DCT coefficients of each segment in the input facial image are compared to their counterparts in general faces. Considering weight of each segment and fuzzy rule sets, the input image is assigned to the closest class. The proposed method is evaluated on FERET and AR databases. Experimental results show that the proposed method is tolerant of variations such as facial expression, occlusion, and rotation.

Keywords: *Gender estimation; Single face image; Image partitioning; Image entropy.*

1. Introduction

Face is one of the most important human's biometrics which is used frequently in every day communication. Face images convey significant amount of information including identity, emotional, age, gender and etc. Extraction of this information from facial images is still an active research field.

Gender estimation based on facial images is a challenging task in pattern recognition with many applications in the fields like person identification and verification, video surveillance, and human-computer interaction.

The previous efforts for gender estimation were based on training processes, using several samples for each person. Some of these methods will be described as literature review in the next section.

Preparing multi-training image samples from different point of views or under different lightening conditions is usually difficult or even impossible in some applications such as video surveillance or criminal cases. Recently, many face recognition and understanding applications have focused on single image per person circumstance [1]. In these non-training methods, first the input image is segmented into sub-images. Then, each of these segments is mapped into appropriate feature space. Finally, with the help of a distance criterion, the extracted feature vector is assigned to the closest class.

In this paper we proposed a new gender estimation algorithm based on single image per person condition. After partitioning the input image into seven segments, Discrete Cosine Transform (DCT) coefficients are obtained from each segment. Since these segments do not have the same importance, the role of each segment is determined according to its entropy. Unlike uniform and trivial segments (including forehead or cheek), those segments such as eyes, eyebrows, nose and mouth with more information (more pixel variation) are highlighted in the entropy image. In the next step, a general face is constructed from male and female images separately. Afterward, feature vectors of the input image and the two general faces are compared and the closest one is chosen as the estimated gender. To improve the gender estimation method, a fuzzy system is implemented. To evaluate the proposed algorithm, we used gray level images in FERET [2] and AR [3] databases [4, 5]. Since women usually have longer hair than men, all images were cropped and resized to 191×164 pixels to contain hair style information. Some samples are shown in Figure 1.

Figure 1. Some Frontal Samples of AR (first row) and FERET (second row)

2. Related Work

As mentioned in previous section, preceding methods for gender estimation were based on training processes which need several samples for each person. Some of these methods are reviewed in this part.

Tivive and Bouzerdoum [6] applied shunting inhibitory convolutional neural networks (SICoNNets) for gender classification [7]. Moghaddam and Yang [8] used a Support Vector Machine (SVM) with Radial Basis Function kernels to classify genders from low resolution 21×12 "thumbnail'' faces. Gutta et al. [9] proposed a hybrid approach that consists of a collection of neural networks and decision trees.

Makinen and Raisamo [10] selected four gender classification methods: a multilayer neural network with pixel-based input, an SVM with pixel-based input, a discrete Adaboost with haar-like features, and an SVM with LBP features. Jain and Huang [11] combined Independent Component Analysis (ICA) and Support Vector Machines (SVMs) for gender classification. Leng and Wang [12] proposed a novel approach based on Fuzzy SVM (FSVM) for gender classification task.

In [13], Sun et al. used principle component analysis (PCA) to compress the input space into low dimensional features (eigen-vectors) and applied a genetic algorithm (GA) for selecting a set of eigen-vectors to represent two classes. Iga et al. [14] have developed an algorithm which estimates gender and age using support vector machine (SVM) classifiers and voting based on features such as geometric arrangement, texture and luminosity patterns extracted from facial images.

J.Wu et al. [15] presented two novel principal geodesic analysis (PGA) methods, weighted PGA and supervised PGA, to parameterize the facial needle-maps, and used the Bayes classifier to classify male and female facial images. Wilhelm et al. [16] compared two models for extracting features from face images. These models were a description of face images by their projection on independent base images, and an Active Appearance Model (AAM) which describes the shape and grey value variations of the face images. They used Nearest Neighbor, MLP, RBF and LVQ networks for classification of extracted feature vectors. Xu et al. [17] presented a hybrid gender classification method by fusing appearance features and geometry features. They used Adaboost algorithm to extract appearance features and combined with geometry features which are extracted with Active Appearance Model. They also used SVM with the Radial Basis Function (RBF) kernel as the classifier.

Golomb et al. [18] trained a fully connected two-layer network, called SEXNET, to identify gender from facial images. Tamura et al. [19] used a multi layered neural network to identify gender from face images of different resolutions. Burton et al. [20] identified 73 points on a face image and discriminate analysis was used to classify gender using point-to-point distances. Brunelli and Poggio [21] compute 16 geometric features (like pupil to eyebrow separation, eyebrow thickness, etc.) from the neutral images of a face. These features were used for identifying the gender. B.Wu et al. [22] employed Adaboost learning algorithm to select a series of look-up table (LUT) weak classifiers for gender classification. Ben Abdelkader and Griffin [23] presented a method that extracted regions from the face and used them as input for an SVM or Fisher Linear Discriminant (FLD) gender classifier. Finally, Takimoto et al. [24] proposed Advanced Retinal Sampling Method (ARSM) as a feature-point detection method. They also used neural networks for sex classification.

3. The Proposed Gender Estimation Algorithm

To be compatible with a single image per person face recognition system, the proposed algorithm uses a non-training procedure for gender estimation. This strategy tackles any inconvenience concerning training data collection or retraining process after new class (person) addition. The main steps of the proposed algorithm are as follows (Figure 2):

1. The input image is segmented.
2. The general faces are created from male and female training images separately.
3. Male and female general faces are also segmented similar to step 1.
4. DCT coefficients are extracted from each segment.
5. Distance between DCT coefficients vector of each segment in the input image and its corresponding segment vector in general faces is calculated.
6. Each segment of input image is initially assigned to the male or female class according to its minimum distance in step 5.
7. The highlighted image is constructed by the input image mapping.

8. The highlighted image is segmented similar to the step1.
9. Entropy of each segment in the highlighted image is calculated. Since the obtained entropy is positive, to differentiate between male and female classes, the entropy of female segments is multiplied by -1.
10. According to the initial estimated gender in step 6, and the entropy of each segment calculated in step 9, a number is assigned to each segment. Absolute of this number indicates the importance (weight) of that segment in the decision process and its sign shows male or female classes.
11. Gender estimation is again estimated based on sum of 7 segments in each image. (Positive sum will be estimated for male and negative for female).
12. Fuzzy system is performed to increase the accuracy.

In the following, details of the proposed gender estimation are described.

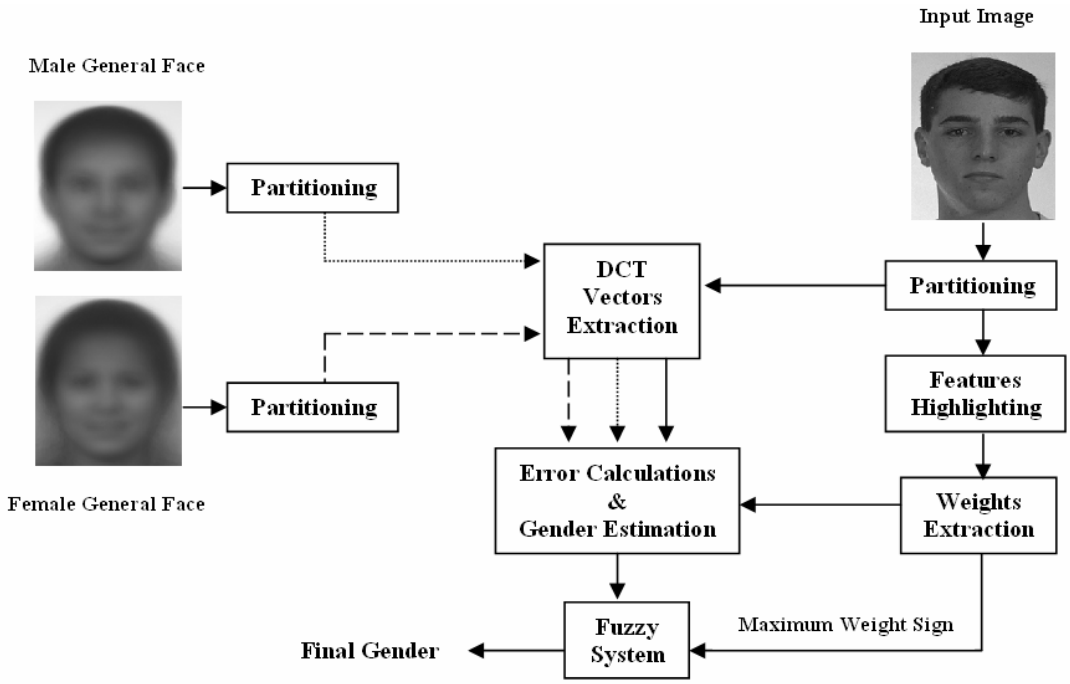

Figure 2. The Framework of Proposed Gender Estimation Algorithm

3.1. Image Datasets

In this paper we use some image samples from FERET and AR datasets [4, 5]. FERET is a general database including numerous facial images obtained under different lightening conditions and facial expressions. Neutral (frontal facial image without any expression) and smile images of 1190 persons in FERET database are utilized in this paper (first and second rows of Figure 5). Then, this database is partitioned into four distinct datasets. FERET Dataset1 (FD1) and FERET Dataset2 (FD2) include neutral images of 348 male and 247 female persons (595 different persons). Smile samples of those persons participated in FD1 and FD2 are classified into FERET Dataset3 (FD3) and FERET Dataset4 (FD4), respectively.

In addition to FERET database, facial images of AR database are also used to evaluate performance of the proposed gender recognition algorithm. AR database consists of 56 female and 70 male persons. Neutral images of these 126 persons are called AR Dataset1 (AR1). AR Dataset2 (AR2) includes smile images of those persons.

3.2. General Face Construction

Since the proposed algorithm for gender estimation is non-training, a reference is needed to classify male and female images which is called *general face* image. General face of a dataset is linear combination of its images as $Gf = \frac{1}{n}(im_1 + im_2 + ... + im_n)$ in which n is number of images in the dataset, im_1 is ith image sample and *Gf* is the general face.

Figure 3 and Figure 4 show male and female general faces of set FD1 as a function of number of male (NM) and female (NF) images participating in this process. Considering these figures, it is obvious that general faces are approximately converged when more than 80 images used in averaging process (NM and NF>80). It is also noticeable that final male and female general faces (those obtained by NM=348 and NF=247) can be differentiated by human observers.

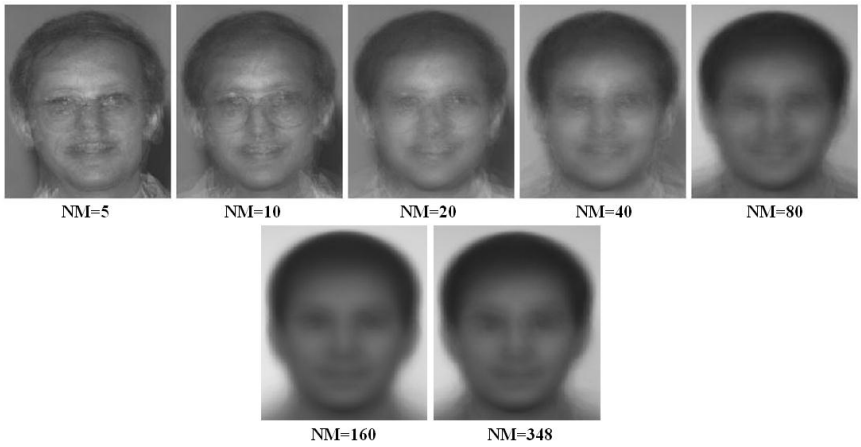

Figure 3. General Face for Different NM in FD1

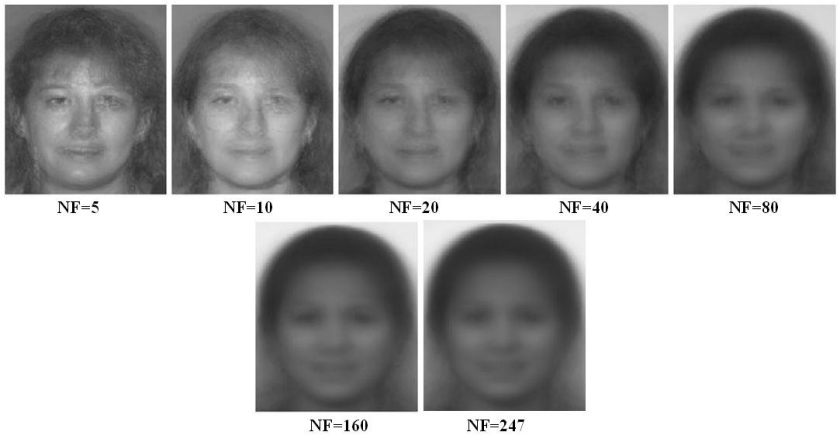

Figure 4. General Face for Different NF in FD1

Regarding Figure 3 (NM=348) and Figure 4 (NF=247), some physiological differences are obvious:

1. The female general face is more rounded
2. The male hairline is usually higher than the female and tends to be more "M" like shape.
3. The male forehead is wider than female.
4. The female eyes are more highlighted.
5. The female cheeks tend to be fuller and more rounded than male ones and the cheekbone itself tends to stand a little bit higher and further forward.
6. The female mouth is usually smaller but more salient.
7. The female chins tend to be rounded while male chins tend to be wider with a flat base and two corners to form a more square shape.
8. The female face background tends to darker because of the long hairs coming down to the chin.

3.3. Image Partitioning

Segmenting into different parts allows extraction of facial features such as eyes, brow, nose and mouth which can differentiate between males and females. In this paper, all images including general faces and test samples are partitioned in the same manner. According to Figure 5, each image is segmented into 6 partitions (4 horizontal and 2 vertical segments), each including *at least* on face feature (eye, eyebrow, nose, mouth). Regarding the original image, 7 partitions are available. Since images have the same size, partitioning can be performed automatically. Then segmented images are manually verified.

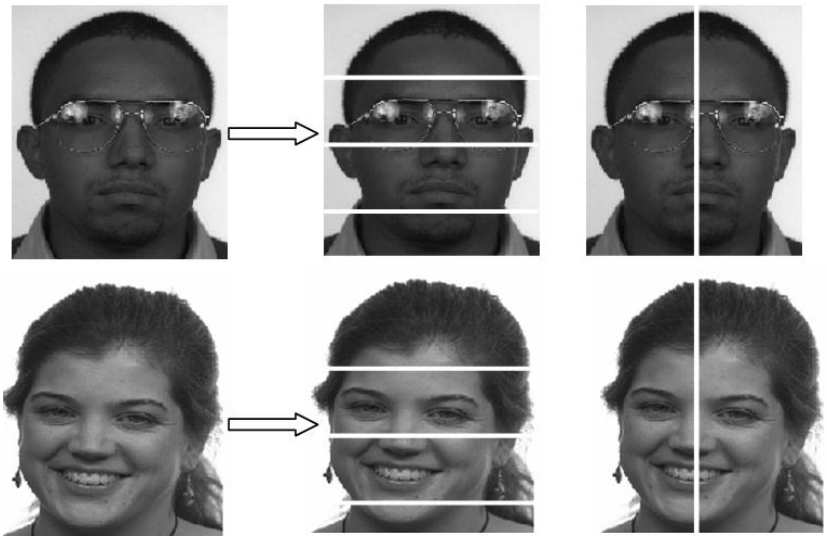

Figure 5. Image Partitioning

3.4. Face Features Highlighting

All parts in a face image do not have the same importance in gender estimation or face recognition. For instance, eyes play more significance role in those scenarios that are occluded in some public images for privacy purpose.

In order to have a better gender estimation, we need to find and highlight those facial parts which are crucial for this task. To highlight face features, we used linear spatial filter, Sobel detectors and Laplacian of Gaussian (LOG) to obtain two point spread function (PSF). Then, the input image is filtered with Sobel and LOG PSF functions [25]. The two output images have different information. When these two images are merged by wavelet decomposition at the first level [26, 27], the output image represents face features in a better way. Those face features like eyes and mouth which are more important for gender estimation are intensified in the highlighted image (Figure 6)

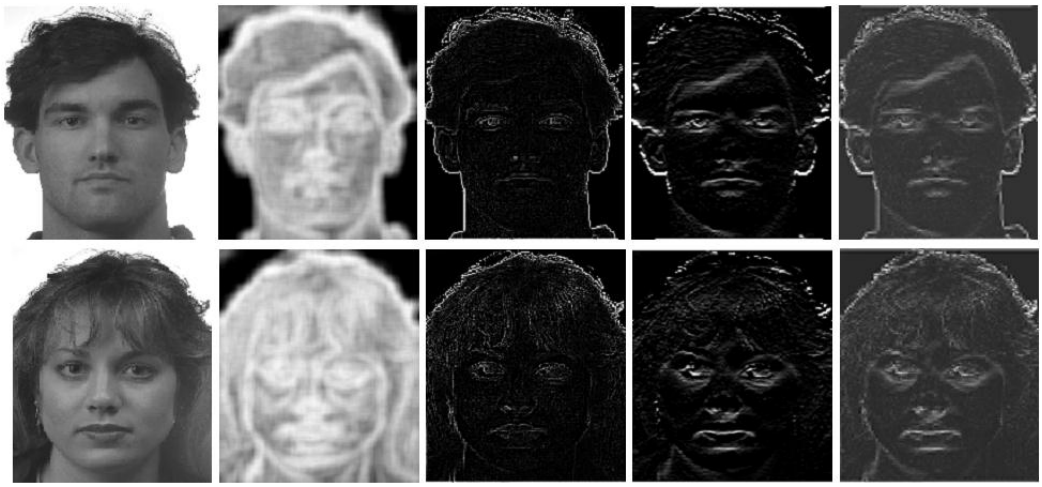

Figure 6. From left to right: the input image and its corresponding entropy image, LOG highlighting, Sobel highlighting and Fusion of LOG and Sobel

3.5. Image Entropy

We used image entropy to determine which partition of the highlighted image has higher weight in the gender estimation process. Those partitions with higher entropy (weight) are considered as more important blocks [28].

Based on Shannon theory [29], the amount of information of a discrete random variable *X(x1,x2,…xs)* with a probability *P(x)(p(x1),p(x2),...,p(x3))* can be described by entropy *H(x)* as follow:

$$H(x) = \sum_{i=1}^{s} p(xi)\log\left(\frac{1}{p(xi)}\right) = -\sum_{i=1}^{s} p(xi)\log(p(xi)) \quad (1)$$

For a digital image f(x,y), entropy can be defined as:

$$H[f(x,y)] = -\sum_{i=1}^{s} pi.\log(pi) \quad (2)$$

Where *pi* is the probability of the i[th] gray level value in the image and *s* is the total number of gray level values. Local entropy in a digital image partition provides a quantitative measure of the information contained in that partition. Normalized entropies of the 7 partitions in Figure 5 are shown in Table 1. The partition including eyes and eyebrows has the highest amount of entropy (information).

Table 1. Amount of Entropy for Partitions of Figure 5

Input Image	*Normalized entropy for each partition*						
	Head	Eye & Brow	Nose & Mouth	Chin	Half Left	Half Right	Total Image
Male	0.920	1	0.796	0.768	0.966	0.909	0.950
Female	0.961	0.970	1	0.867	0.999	0.932	0.981

3.6. Discrete Cosine Transform

Discrete Cosine Transform (DCT) is a well-known transformation technique used for image compression applications. Our idea is to use the most significant DCT coefficients to represent the image space, and classify each image based on these coefficients [30, 31]. DCT transform for a 2D image, *f(i,j)*, with *N* columns and *M* rows is defined as[32, 33]:

$$F(U,V)=\left(\frac{2}{N}\right)^{\frac{1}{2}}\left(\frac{2}{M}\right)^{\frac{1}{2}}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}A(i)A(j).\cos\left[\frac{\pi.u}{2.N}(2i+1)\right].\cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j) \quad (3)$$

In which *A(i)* is:

$$A(i)=\begin{cases}\frac{1}{\sqrt{2}} & for \quad i=0\\ 1 & otherwise\end{cases} \quad (4)$$

The most significant DCT coefficients are chosen from the given face image by determining which coefficients have the greatest variance. In natural images, the DCT coefficients variances drop for higher frequency components. Therefore, we use 10×10 DCT coefficients matrix from low frequency space for gender estimation.

3.7. Distance Measures

Some different distance criteria have been utilized in this paper to measure the distance between partitions in the input image and their corresponding in the general faces. Let *X*, *Y* be vectors of length *n*. Then we can calculate the following distances between these vectors [34].

- Minkowski distance (L_p Metrics): $d(X,Y)=L_p(X,Y)=\left(\sum_{i=1}^{n}|x_i-y_i|^p\right)^{\frac{1}{p}}$ (5)

  Manhattan distance (L_1 Metrics, city block distance):

  $d(X,Y)=L_{p=1}(X,Y)=\sum_{i=1}^{n}|x_i-y_i|;$ (6)

- Euclidean distance (L_2 Metrics): $d(X,Y)=L_{p=2}(X,Y)=\|X-Y\|=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2};$ (7)

- Squared Euclidean distance (sum square error, SSE), mean square error (MSE):

$$d(X,Y) = L_{p=2}^2(X,Y) = SSE = \|X - Y\|^2 = \sum_{i=1}^{n}(x_i - y_i)^2; \quad (8)$$

$$d(X,Y) = \frac{1}{n} L_{p=2}^2(X,Y) = MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2; \quad (9)$$

- Cosine distance: $$d(X,Y) = -\cos(X,Y) = -\frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2}}; \quad (10)$$

- Correlation distance: $$d(X,Y) = -\frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)(n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2)}}; \quad (11)$$

- Chi square distance: $$d(X,Y) = \sum_{i=1}^{n}\frac{(x_i - y_i)^2}{x_i + y_i}; \quad (12)$$

- Canberra distance: $$d(X,Y) = \sum_{i=1}^{n}\frac{|x_i - y_i|}{|x_i| + |y_i|}; \quad (13)$$

- Modified Manhattan distance: $$d(X,Y) = \frac{\sum_{i=1}^{n}|x_i - y_i|}{\sum_{i=1}^{n}|x_i|\sum_{i=1}^{n}|y_i|}; \quad (14)$$

- Modified SSE-based distance: $$d(X,Y) = \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2}; \quad (15)$$

- Chebyshev distance: $$d(X,Y) = \max_i(|x_i - y_i|); \quad (16)$$

  This equals the limit of the L_p Metrics: $\lim_{p\to\infty}(\sum_{i=1}^{n}|x_i - y_i|^p)^{\frac{1}{p}}$;

- Spearman's rank correlation: $$d(X,Y) = \frac{n(\sum_{i=1}^{n} x_i y_i) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2}\sqrt{n(\sum_{i=1}^{n} y_i^2) - (\sum_{i=1}^{n} y_i)^2}}; \quad (17)$$

3.8. Fuzzy System

In this paper fuzzy logic is used to increase accuracy of the proposed gender estimation method. The fuzzy system has two inputs, one from total weights of 7 partitions (*tot_wei*), and the other from maximum weight sign (*max_wei_sign*) of 7 partitions.

In some cases, sum of total weights of 7 partitions is around zero (*-0.7<w<0.7*), indicating an error case. To determine the final gender of the input image (*final_tot_wei*), the fuzzy

system utilizes total weight and the maximum weight sign according to heuristic fuzzy rules and predefined thresholds. Membership function is Gaussian curve built-in type.

The rules are as follows:

1) If (*tot_wei* is *LN*) and (*max_wei_sign* is *Negative*) then (*final_tot_wei* is *LN*)
2) If (*tot_wei* is *LN*) and (*max_wei_sign* is *Positive*) then (*final_tot_wei* is *LP*)
3) If (*tot_wei* is *LP*) and (*max_wei_sign* is *Negative*) then (*final_tot_wei* is *LN*)
4) If (*tot_wei* is *LP*) and (*max_wei_sign* is *Positive*) then (*final_tot_wei* is *LP*)
5) If (*tot_wei* is *HN*) and (*max_wei_sign* is *Negative*) then (*final_tot_wei* is *HN*)
6) If (*tot_wei* is *HN*) and (*max_wei_sign* is *Positive*) then (*final_tot_wei* is *HN*)
7) If (*tot_wei* is *HP*) and (*max_wei_sign* is *Negative*) then (*final_tot_wei* is *HP*)
8) If (*tot_wei* is *HP*) and (*max_wei_sign* is *Positive*) then (*final_tot_wei* is *HP*)

LN:-0.7<W<0; HN: W<-0.7; LP: 0<W<0.7; HP: W>0.7

4. Experimental Results

As explained in section 3.1, we used facial images of 1190 persons in FERET database and 126 persons in AR database, called FD1, FD2, FD3, FD4, AD1, and AD2. These datasets are distinct and contain neutral and smile facial images. This section explores robustness of the proposed gender estimation algorithm under several circumstances such as: different distance measures, different number of images participated in general face construction, effect of fuzzy system, head rotation and occlusion. It should be mentioned that in the following experiments, general faces are constructed from neutral images, while test samples are chosen from dataset including smile facial images (FD3, FD4, and AD2).

4.1. Distance Measures Study

Different distance measures were introduced for single image face recognition discussed in section 3.7. In the first experiment we study accuracy of these distance measures for gender estimation. Male and female general faces are constructed from neutral images in FD1 with NM=348 and NF=247. Test samples are smile facial images in FD3. As Table 2 shows, Modified SSE has the highest performance over the entire test dataset. Similar experiment with AR database indicates that Modified SSE distance outperforms the others with 89.68% accuracy. In the rest of this paper, we will use Modified SSE distance measure for gender estimation.

Table 2. Estimation Rate of Various Distance Measures

Distance measures	Test Images		
	Males of FD3 (348)	Females of FD3 (247)	Total (595)
Minkowski	82.75%	76.51%	80.16%
Manhattan	82.75%	78.13%	80.84%
Euclidean	82.75%	76.51%	80.16%
Squared Euclidean distance	82.75%	76.51%	80.16%
mean square error	82.75%	76.51%	80.16%
Cosine	81.89%	75.30%	79.15%
Correlation	81.89%	75.30%	79.15%
Chi square	60.34%	57.08%	58.99%
Canberra	82.75%	74.08%	79.15%
Modified Manhattan	94.25%	60.72%	80.33%
Modified SSE	87.93%	74.08%	82.18%
Chebyshev	82.75%	74.49%	79.32%
Spearman's rank correlation	78.44%	80.56%	79.32%

4.2. Number of Images Effect

Using more facial images for general face construction, higher accuracy in gender estimation can be obtained. In the second experiment, we consider performance of the proposed method as a function of male and female image numbers (NM and NF).

Figure 7 shows ability of the proposed gender estimation system which uses images of FD1 for general face construction, and smile samples of FD3 for performance evaluation. Likewise, in Figure 8, general faces are constructed from AD1 while AD2 is used for testing. In Figure 9, dataset FD1 is used for both general faces creation and evaluation. The same experiment is performed with dataset AD1 (Figure 10).

As Figure 7 and Figure 8 show, accuracy of the gender estimation system is increased when more samples are used for general face construction. Therefore, next experiments will be based on maximum number of male and female images (NM=348 and NF=247). Comparing Figure 7 and Figure 8 with their counterparts in Figure 9 and Figure 10 indicate robustness of the proposed method. Using the same datasets for general faces construction and testing slightly outperforms the case in which datasets are different. It is noticeable that the gender estimation is more accurate for men than women. It is mostly due to hair variations in female images.

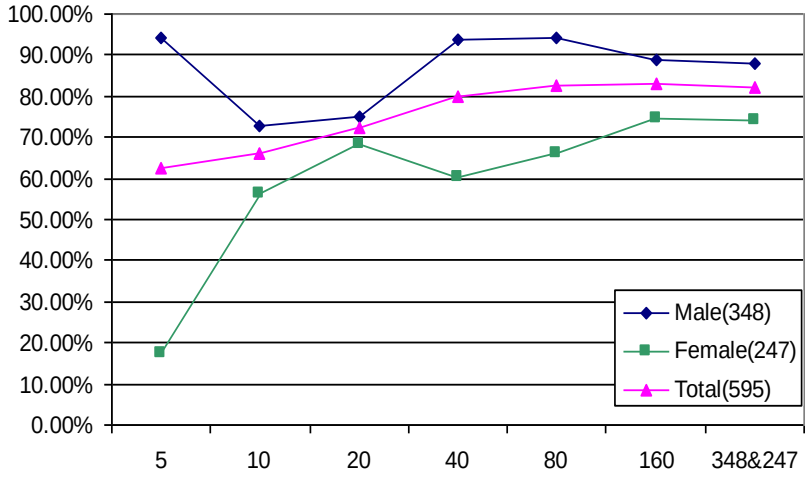

Figure 7. Estimation Rates for Different NM and NF using FD3 Test Images

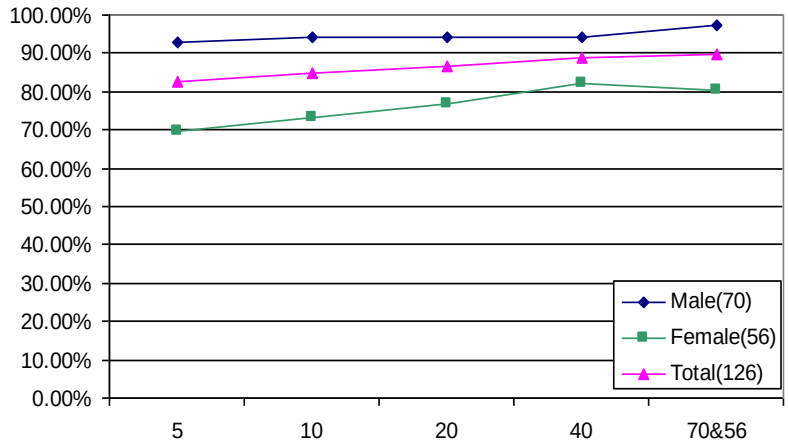

Figure 8. Estimation Rates for Different NM and NF using AD2 Test Images

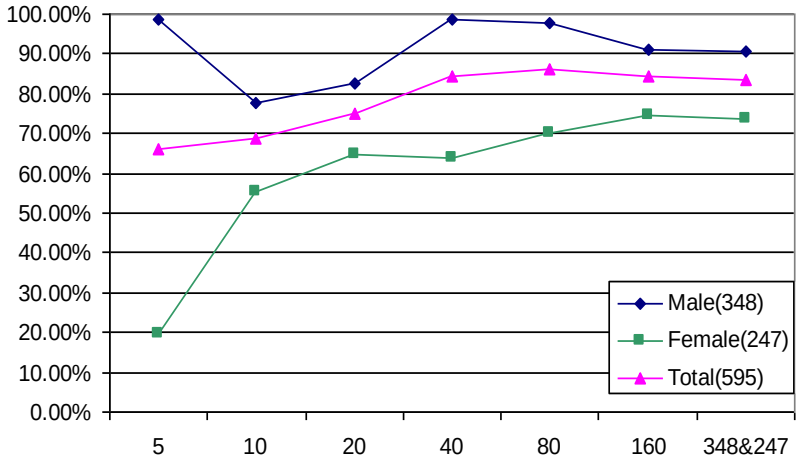

Figure 9. Estimation Rates for Different NM and NF using FD1 Test Images

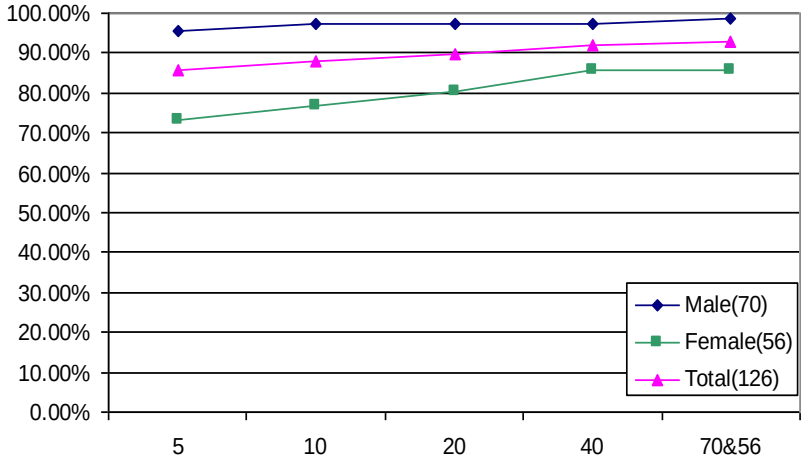

Figure 10. Estimation Rates for Different NM and NF using AD1 Test Images

4.3. Fuzzy Logic Effect

As mentioned before, fuzzy logic system is used to augment accuracy of the gender estimation method. Table 3 shows that fuzzy logic system considerably enhanced the accuracy. To evaluate the proposed system with test samples in FD3 and FD4, general faces are constructed from FD1 and FD2 respectively. AD1 is used for general face construction when AD2 is used for testing.

Table 3. Fuzzy Effect

Test Images	Without Fuzzy	With Fuzzy	Enhancement
Males of FD3 (348)	82.18%	87.93%	5.75%
Females of FD3 (247)	67.20%	74.08%	6.88%
FD3 (595)	75.96%	82.18%	6.22%
Males of FD4 (348)	88.21%	92.52%	4.31%
Females of FD4 (247)	56.27%	63.15%	6.88%
FD4 (595)	74.95%	80.33%	5.38%
Males of AD2 (70)	94.28%	97.14%	2.86%
Females of AD2 (56)	76.78%	80.35%	3.57%
AD2 (126)	86.50%	89.68%	3.18%

4.4. Performance Evaluation Using Different Datasets

Generalizing power of the proposed algorithm is studied in this experiment. Here, general faces are constructed from one database (FERET or AR), while images of another databases are used for evaluation. Like other experiments, neutral images are used for general face creation and smile images are used for testing.

As Table 4 shows, when general faces are produced from FERET database and the algorithm is tested by AR dataset or vice versa, accuracy of the gender estimation algorithm is slightly decreased in the comparison with the case when one database is used for both general face construction and testing. This means, the proposed gender estimation algorithm is not sensitive to imaging condition such as illumination and background as it is different in FERET and AR.

Table 4. Estimation Rates for Various General Faces

	General Faces		
Test Images	FD1	FD2	AD1
Males of FD3 (348)	87.93%	91.66%	83.04%
Females of FD3 (247)	74.08%	66.80%	74.49%
FD3 (595)	82.18%	81.34%	79.49%
Males of FD4 (348)	91.37%	92.52%	88.21%
Females of FD4 (247)	59.91%	63.15%	64.37%
FD4 (595)	78.31%	80.33%	78.31%
Males of AD2 (70)	68.57%	88.57%	97.14%
Females of AD2 (56)	76.78%	71.42%	80.35%
AD2 (126)	72.22%	80.95%	89.68%

4.5. Performance Evaluation Using Non-frontal Facial Images

In addition to frontal samples, FERET database also contains non-frontal facial images such as half right view and profile samples for each person. As another test case, 80 persons are randomly selected from FERET database. Then, right view and profile samples of each person are used in this experiment. The gender estimation algorithm uses general faces constructed from frontal samples to classify input images to male and female categories. Non-

frontal facial images are resized and partitioned in the same manner as frontal samples discussed in the previous sections. Figure 11 shows some samples of right view and profile images.

Figure 11. Non-frontal samples of FERET: Half right view (first row) and profile (second row)

Table 5. Estimation Rates of Non-frontal Images

	General Faces			
Test Images	**FD1**	**Half right view**	**Profile**	**AD1**
Males of FD3	87.93%	90.80%	-	83.04%
Females of FD3	74.08%	54.65%	-	74.49%
FD3 (Total)	82.18%	75.79%	-	79.49%
Males of Half right view	85.00%	93.75%	88.75%	72.50%
Females of Half right view	45.00%	71.25%	57.50%	63.75%
Half right view (Total)	65.00%	82.50%	73.12%	68.12%
Males of Profile	-	78.75%	82.50%	-
Females of Profile	-	75.00%	81.25%	-
Profile (Total)	-	76.87%	81.87%	-
Males of AD2	68.57%	75.71%	-	97.14%
Females of AD2	76.78%	32.14%	-	80.35%
AD2 (Total)	72.22%	56.34%	-	89.68%

As Table 5 shows, maximum accuracy is obtained in the cases in which the same dataset is used for both general face construction and gender estimation. Moreover, to have an acceptable precision, difference between head positions in images used for general face construction and gender estimation should be less than 90 degree. In the other words, it is not applicable to construct the general face from neutral images and use profile facial images for testing or vice versa. However, the proposed system is able to classify male and female images with half view images.

4.6. Performance Evaluation with Occulted Images

In this experiment, robustness of the proposed gender estimation algorithm is explored under occlusion. 70 images of male and 56 images of female persons are occluded in AR database with glass and scarf (Figure 12). As Table 6 shows, accuracy of the proposed gender estimation algorithm is declined when facial features are occluded with glass or scarf. However, gender estimation accuracy is increased for females when their faces are occluded with scarf. The lowest horizontal segment in general faces in Figure 3 (NM=348) and Figure 4 (NF=247) is darker for females. Therefore, the scarf increases role of this segment and consequently most of occluded input images are assigned to female class.

Figure 12. Samples for Glass and Scarf Occlusion

Table 6. Estimation Rates for Occlusion

Test images	Occlusion type		
	Glass	Scarf	Without occlusion
Males of AR	85.71%	62.85%	97.14%
Females of AR	71.42%	92.85%	80.35%
Total AR	79.36%	76.19%	89.68%

4.7. Performance Evaluation Using an Ethnic Dataset

As another criterion for generalization power of the proposed method, in the last experiment we used an unconstrained ethnic dataset gathered trough Internet. In this dataset all women have scarf and most of met have beard and moustache. To be tangible when we compare the obtained results on ethnic dataset with AR dataset, the new dataset also includes facial images of 70 males and 56 females (like AR dataset). Each person in ethnic dataset has only one frontal image, which means test samples are identical with those participating in general image construction process. Figures 13 and 14 show some samples and general images in ethnic database. According to Figure 14, effect of scarf and moustache is noticeable in female and male general images.

Table 7 compares accuracy of the proposed method on ethnic and AR datasets. It can be concluded that effect of scarf in females of ethnic dataset considerably increases the accuracy. It is mainly due to hair variation removal by scarf. Unlike previous experiments in which algorithm's accuracy for female were less than males, scarf increases the accuracy for females. Female and male general images of ethnic dataset are very different from each other compared to their counterparts in AR. This led to higher overall accuracy for ethnic dataset.

Figure 13. Female (first row) and Male (second row) Samples from Ethnic Dataset

Figure 14. Male (right) and Female (left) General Faces of Ethnic Dataset

Table 7. Estimation Rates for Occlusion

Ethnic Test images	*General Faces*		
	Ethnic	FD1	AD1
Male (70)	90.00%	81.42%	84.28%
Female (56)	94.64%	73.21%	78.57%
Total (126)	92.06%	77.77%	81.74%

4.8. Comparison with Related Works

So far, several methods have been proposed for gender estimation which is based on training procedures. Table 8 tabulates those methods which used FERET and AR databases. Because of unavailability of other single image gender estimation methods to the authors, it is very difficult to compare the obtained results with them.

According to Table 8, the obtained results are comparable with accuracy of training based gender estimation methods that used small parts of FERET database [10]. It should be mentioned that in the systems which used larger parts of FERET [7, 8, 9], number of persons are not clearly stated. Considering that previous efforts for gender estimation use multiple training samples for each person, the proposed method yields promising results.

Table 8. Different Methods for Gender Estimation

Method	Database	# Male and Female images		Estimation rate (Max)	Reference
SICoNNets	FERET	1152	610	97.1 %	[7]
SVM *with* Gaussian RBF kernel	FERET	1044	711	96.62 %	[8]
SVM *with* Cubic polynomial kernel	FERET	1044	711	95.12 %	[8]
ERBF + DT	FERET	1906	1100	96 %	[9]
SVM with LBP features	FERET	212	199	79.17 %	[10]
Neural Network with face pixels	FERET	212	199	83.30 %	[10]
SVM with face pixels	FERET	212	199	83.38 %	[10]
Adaboost with haar-like features	FERET	212	199	82.60 %	[10]
ICA + COS	FERET	250	250	85.33 %	[11]
ICA + LDA	FERET	250	250	93.33 %	[11]
ICA + SVM	FERET	250	250	95.67 %	[11]
Fuzzy SVM *with* Gabor features	FERET	160	140	98.09 %	[12]
SVM *with* Gabor features	FERET	160	140	96.5 %	[12]
LDA *with* Gabor features	FERET	160	140	93.33 %	[12]
NN *with* Gabor features	FERET	160	140	91.5 %	[12]
Supervised PGA with Bayes classifier	AR	74	94	97.5%	[15]
AAM & Geometry features with SVM	FERET & AR	500	500	92.38%	[17]
Proposed method	FERET	696	494	82.18%	
Proposed method	AR	70	56	89.68%	

5. Conclusion and Future Work

In this paper a novel gender estimation algorithm was proposed. Unlike previous efforts which try to estimate the gender of the input image via a training process, our algorithm was based on single image per person. The proposed gender estimation algorithm was evaluated

on FERET and AR databases and maximum accuracy of 82.18% and 89.68% was obtained respectively. Different experimental results such as face rotation and occlusion showed robustness of the proposed method.

Considering FERET database as a large and diverse image database, the experimental results can be promising that encourage us to study the effect of the proposed gender estimation algorithm on a face recognizer. For the future work we will use the proposed gender estimation algorithm prior to the face recognition. This strategy will prune the overall database into male and female parts and as a result, the recognition time will be decreased. If the gender estimation error be less that recognizer misclassification error, we can hope to enhance recognition time and accuracy at the same time.

Acknowledgements

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

References

[1] Tan, X., Chen, S., Zhou, Z.H., Zhang, F., Face recognition from a single image per person: a survey. Pattern Recognition, vol. 39, **(2006)**, pp. 1725 – 1745.

[2] Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J., The FERET Evaluation Methodology for Face Recognition Algorithms. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, **(2000)**, pp. 1090-1104.

[3] Martiniz, A.M., Benavente, R., The AR face database, Technical Report, vol. 24, CVC, **(1998)**.

[4] http://face.nist.gov/colorferet/request.html.

[5] http://rvl1.ecn.purdue.edu/~aleix/ar.html.

[6] Tivive, F.H.C., Bouzerdoum, A., A shunting inhibitory convolutional neural network for gender classification, International Conference on Pattern Recognition, **(2006)**.

[7] Tivive, F.H.C., Bouzerdoum, A., Efficient training algorithms for a class of shunting inhibitory convolutional neural network, IEEE Transactions on Neural Networks, vol. 16(3), **(2005)**, pp. 541-556.

[8] Moghaddam, B., Yang, M.H., Gender Classification with Support Vector Machines, IEEE International Conference on Automatic Face and Gesture Recognition, **(2000)**, pp. 306-311.

[9] Gutta, S., Weschler, H., Phillips, P.J., Gender and Ethnic Classification of Human Face using Hybrid classifier. IEEE International Conf. on Automatic Face and Gesture Recognition, **(1998)**, pp. 194-199.

[10] Makinen, E., Raisamo, R., Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, **(2008)**, pp. 541-547.

[11] Jain, A., Huang, J., Integrating independent components and support vector machines for gender classification, International Conference on Pattern Recognition, **(2004)**.

[12] Leng, X.M., Wang, Y.D., Improving generalization for gender classification, IEEE International Conference on Image Processing, **(2008)**, pp. 1656-1659.

[13] Sun, Z., Yuan, X., Bebis, G., Louis, S., Neural network-based gender classification using genetic eigen-feature extraction, International joint conference on neural networks, vol. 3, **(2002)**, pp. 2433-2438.

[14] Iga, R., Izumi, K., Hayashi, H., Fukano, G., Ohtani, T., A Gender and Age Estimation System from Face Images, SICE Annual Conference in Fukui, **(2003)**, pp. 756-761.

[15] Wu, J., Smith, W.A.P., Hancock, E.R., Gender Classification based on Facial Surface Normals, International Conference on Pattern Recognition, **(2008)**.

[16] Wilhelm, T., Bohme, H.J., Gross, H.M., Classification of face images for gender, age, facial expression and identity, International Conference on Artificial Neural Networks, **(2005)**, pp. 569-574.

[17] Xu, Z., Lu, L., Shi, P., A Hybrid Approach to Gender Classification from Face Images, International Conference on Pattern Recognition, **(2008)**.

[18] Golomb, B., Lawrence, D., Sejnowski, T., SEXNET: A neural network identifies sex from human faces, Advances in Neural Information Processing Systems, vol. 3, **(1991)**, pp. 572-577.

[19] Tamura, S.H., Mitsumoto, H., Male/Female identification from 8×6 very low resolution face images by neural network, Pattern Recognition, vol. 29, **(1996)**, pp. 331-335.

[20] Burton, A., Bruce V., Dench, N., What's the Difference Between Men and Women? Evidence from Facial Measurements, Perception, vol. 22, **(1993)**, pp. 153-176.

[21] Poggio, T., Hyberbf Networks for Gender Classification, DARPA Image Understanding Workshop, **(1992)**, pp. 311-314.

[22] Wu, B., Ai, H., Huang, C., LUT-based adaboost for gender classification, International conference on audio and video-based biometry person authentication, **(2003)**, pp. 104-110.

[23] Ben Abdelkader, C., Griffin, P., A Local Region-Based Approach to Gender Classification from Face Images, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, **(2005)**.

[24] Takimoto, H., Mitsukura, Y., Fukumi, M., Akamatsu, N., A Design of Gender and Age Estimation System Based on Facial Knowledge, SICE-ICASE International Joint Conference, **(2006)**, pp. 3883-3886.

[25] Gonzalez, R.C., Woods, R.E., Digital Image Processing, Prentice Hall, Upper Saddle River, NJ, **(2002)**.

[26] Pajares, G., de la Cruz, J.M., A wavelet-based image fusion tutorial, Pattern Recognition, vol. 37, **(2004)**, pp. 1855-1872.

[27] Amolins, K., Zhang, Y., Dare, P., Wavelet based image fusion techniques-An introduction, review and comparison, ISPRS Journal of Photogrammetry & Remote Sensing, vol. 62, **(2007)**, pp. 249-263.

[28] Kanan, H.R., Faez, K., Gao, Y., Face recognition using adaptively weighted patch PZM array from a single exemplar image per person, Pattern Recognition, vol. 41, **(2008)**, pp. 3799 – 3812.

[29] Shannon, C.E., A mathematical theory of communication. Bell Syst. Tech. J., vol. 27, **(1948)**, pp. 379-423, pp. 623-656.

[30] Pan, Z., Adams, R., Bolouri, H., Image Recognition using Discrete Cosine Transformation as Dimensionality Reduction, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, **(2001)**.

[31] Pan, Z., Adams, R., Bolouri, H., Dimensionality Reduction for face Images Using Discrete Cosine Transformation for Recognition, Technical Report, Science and Technology Research Centre, University of Herefordshire, **(2000)**.

[32] Hafed, Z.M., Levine, M.D., Face Recognition Using the Discrete Cosine Transform, International Journal of Computer Vision, vol. 43(3), **(2001)**, pp. 167–188.

[33] Duh, D.J., Jeng, J.H., Chen, S.Y., DCT based simple classification scheme for fractal image compression. Image and Vision Computing, vol. 23, Issue 13, **(2005)**, November, pp. 1115-1121.

[34] Perlibakas, V., Distance measures for PCA-based face recognition, Pattern Recognition Letters, vol. 25, **(2004)**, pp. 711-724.

Authors

Rohollah Akbari received his BSc (2004) and MSc (2009) in electronic engineering from Azad university of Qazvin, Iran. His research interests include Digital Image Processing, Computer Vision and Pattern Recognition.

Saeed Mozaffari received his B.Sc. and M.Sc. degrees in Electrical Engineering from Amirkabir University of Technology, Theran, Iran in 1999 and 2001, respectively. In 2006, he attended Technical University of Braunschweig in Germany as sabbatical leave to finalize his PhD thesis. After obtaining his PhD from Amirkabir University of Technology in 2007, Dr. Mozaffari joined Semnan University in Iran.

His research interests include Digital Image Processing, Computer Vision, Pattern Recognition, and Document analysis.