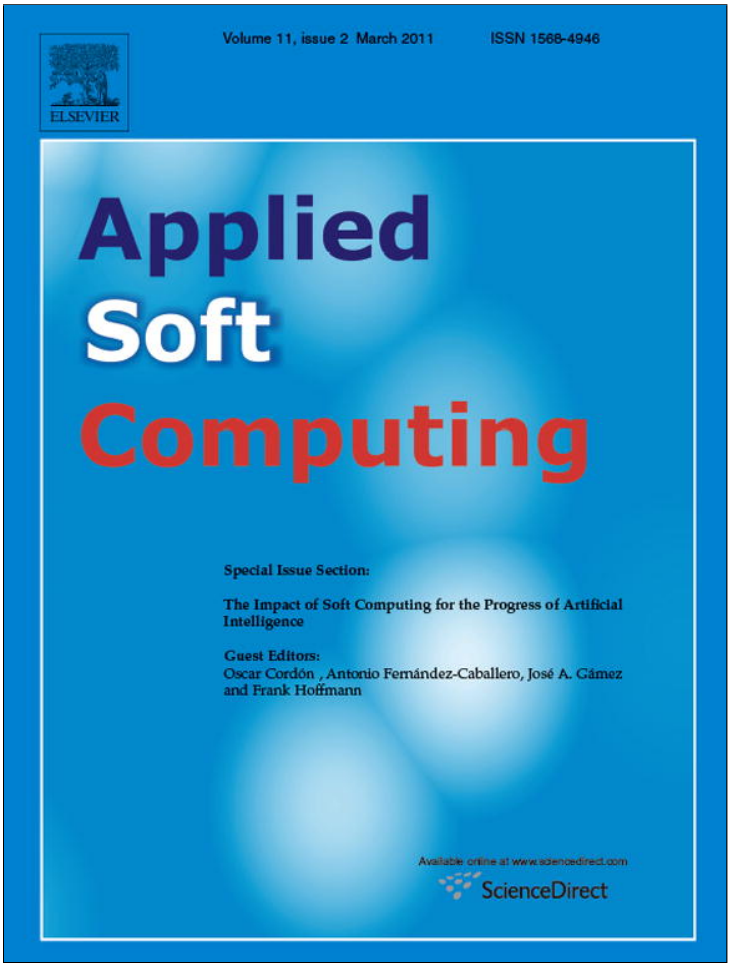
Volume 11, issue 2 March 2011
ISSN 1568-4946
ELSEVIER
Applied
Soft
Computing
Special Issue Section:
The Impact of Soft Computing for the Progress of Artificial Intelligence
Guest Editors:
Oscar Cordón , Antonio Fernández-Caballero, José A. Gámez and Frank Hoffmann
Available online at www.sciencedirect.com
ScienceDirect

Applied Soft Computing 11 (2011) 1801–1810

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A novel fuzzy rule base system for pose independent faces detection

Payman Moallem [a,*], Bibi Somayeh Mousavi [b], S. Amirhassan Monadjemi [c]

[a] Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, HezarJerib Street, Isfahan, Iran
[b] Department of Computer Engineering, Islamic-Azad University, and Member of Young Researchers Club, Zahedan Branch, Zahedan, Iran
[c] Department of Computer Engineering, Faculty of Engineering, University of Isfahan, HezarJerib Street, Isfahan, Iran

A R T I C L E I N F O

Article history:
Received 14 July 2009
Received in revised form 7 May 2010
Accepted 30 May 2010
Available online 8 June 2010

Keywords:
Face detection
Lips detection
Color space
Fuzzy rule-based system
Subtractive clustering
Geometric moment

A B S T R A C T

Human face detection plays an important role in a wide range of applications such as face recognition, surveillance systems, video tracking applications, and image database management. In this paper, a novel fuzzy rule-based system for pose, size, and position independent face detection in color images is proposed. Subtractive clustering method is also applied to decide on the numbers of membership functions. In the proposed system, skin-color, lips position, face shape information and ear texture properties are the key parameters fed to the fuzzy rule-based classifier to extract face candidate in an image. Furthermore, the applied threshold on the face candidates is optimized by genetic algorithm. The proposed system consists of two main stages: the frontal/near frontal face detections and the profile face detection. In the first stage, skin and lips regions are identified in HSI color space, using fuzzy schemes, where the distances of each pixel color to skin-color and lips-color clusters are applied as the input and skin-likelihood and lips-like images are produced as the output. Then, the labeled skin and lips regions are presented to both frontal and profile face detection algorithms. A fuzzy rule-based containing the face and lips position data, along with the lips area and face shape are employed to extract the frontal/near frontal face regions. On the other hand, the profile face detection algorithm uses a geometric moments-based ear texture classification to verify its outcomes. The proposed method is tried on various databases, including HHI, Champion, Caltech, Bao, Essex and IMM databases. It shows about 98, 96 and 90% correct detection rates over 783 samples, in frontal, near frontal and profile face images, respectively.

1. Introduction

Automatic detection of human faces has been the topic of many researches for decades. Human face detection is also becoming a very important part of many applications, such as face recognition, video surveillance and security control systems, intelligent human–computer interface, content-based image retrieval, multimedia applications on web like video conferencing, and face database management.

The goal of the face detection algorithm is to identify the location of all human faces in a given image. Although this issue is so easy for a man brain, it still remains a challenging and difficult problem for a computer. Difficulties such as differences between various facial expressions, image orientation, presence or absence of structural components (e.g. bread, mustaches and glasses), imaging conditions such as lighting (e.g. spectra, source distribution and intensity), camera characteristics (e.g. sensor and lens response) can be mentioned. Again, occlusion, which means that some faces may be partially occluded by other objects in an image with a group of people, can be a serious problem.

Face detection algorithms utilize some particular techniques such as neural networks, support vector machines, hidden Markov model, Adaboost, Hough transform, template matching, principle component analysis, color, texture and facial features analysis to get through an acceptable performance [1]. Each of the mentioned techniques has its own advantages and disadvantages. For instance, neural networks, which have been applied in many pattern recognition and autonomous robot driving problems, can be trained easily to capture the complex class conditional density of face pattern; still, network architecture has to be extensively tuned to achieve an exceptional performance and it is designed to perform a particular task, e.g. to locate only frontal faces in gray scale images [2].

Another technique is skin-color analysis. In color images, skin-color segmentation is one the key feature in many face detection algorithms. Typically when skin-color is utilized as a feature in face detection, two main points should be considered: choosing color space and, modeling of the skin-color distribution.

The choice of color space is considered as the primary step in skin-color segmentation. The RGB color space is the default color space for most available imaging system. Any other color spaces

* Corresponding author.
E-mail addresses: p_moallem@eng.ui.ac.ir, p_moallem@yahoo.com (P. Moallem), bbmoosavi@gmail.com (B.S. Mousavi), monadjemi@eng.ui.ac.ir (S.A. Monadjemi).

1568-4946/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2010.05.024

could be calculated by a linear or non-linear, one-to-one transformation from RGB and vice versa. Several color spaces have been proposed for skin detection. The most widely used color spaces are classified as [3]:

- Basic color spaces (RGB [4], normalised RGB [5], CIE-XYZ).
- Perceptual color spaces (HSI, HSV [6], HSL, TSL).
- Orthogonal color spaces (YCbCr [7], YIQ, YUV, YEs).
- Perceptually uniform color spaces (CIE-Lab and CIE-Luv).
- Others (Mixture spaces, color ratio spaces).

After choosing a suitable color space, the decision rule to discriminate between skin and non-skin pixels should be carried out. To achieve that goal, four groups of skin modeling methods are utilized [8] including: explicitly defined skin regions, parametric and nonparametric skin distribution modeling, and dynamic skin distribution models. Although, color information is a useful tool to identify facial area, skin-color models are not effective where the spectrum of the light source varies significantly [3].

Hsu et al. [7] proposed a face detection algorithm based on a lighting compensation technique and a non-linear color transformation that can be applied in a wide range of the skin-color. Unfortunately, this algorithm could not detect faces with all kind of poses. In fact, many of work in face detection domain are dedicated to special kind pose. However, Jing et al. [9] proposed a system to detect face under various environment and poses, which combines the shape, color, and lighting distribution information.

In a real environment, the face detection problem involves two main challenges including: variable conditions and diversity of human faces; as well as uncertainty of the basic face features such as, skin and lips-color, and face shape. In this paper, a novel pose invariant face detection system is developed, using several accurately designed Fuzzy Inference Systems (FIS) which is supposed to successfully overcome the complexity of the face detection problem. The proposed system contains four FIS's which try to utilize the most reliable features: color information for skin and lips region segmentation; face shape, area and relative position of detected lips region relative to face region for frontal/near frontal faces detection; and finally ear texture analysis for profile face detection.

The remainder of the paper is organized as follows. The next section presents the applied tools and techniques including fuzzy rule-based system and geometric moments. Then Section 3 describes the proposed system including: the applied skin and lips-color segmentation, and the detection algorithms used to detect frontal/near frontal and profile faces. Finally some implementation results on various databases including frontal, near frontal and profile face images are presented and compared to three known robust algorithms to show the effectiveness of the proposed scheme.

2. Applied tools and techniques

2.1. Fuzzy rule-based system

Fuzzy sets theory provides a framework to materialize the fuzzy rule-based (or inference) systems which have been applied to many disciplines such as control systems, decision-making and pattern recognition [10]. The fuzzy rule-based system consists of a fuzzification interface, a rule base, a database, a decision-making unit, and finally a defuzzification interface [11]. These five functional blocks are depicted in Fig. 1 where the rule base and the database are jointly referred to as the knowledge base.

The function of each block is as follow:

- A rule base containing a number of fuzzy IF-THEN rules.

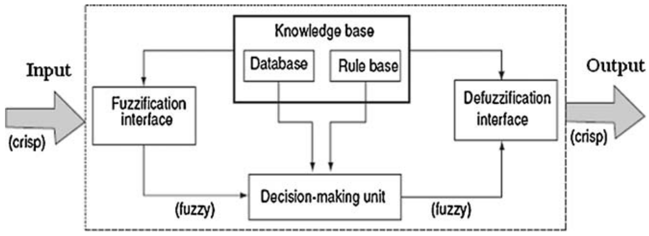

Fig. 1. A typical fuzzy inference system (FIS).

- A database which defines the membership functions (MF) of the fuzzy sets.
- A fuzzification interface which transforms the input crisp values to input fuzzy values.
- A decision-making unit which performs the inference operation on the rules and producing the fuzzy results.
- A defuzzification interface which transform the fuzzy results of the decision-making unit to the crisp output value.

In order to perform the inference operation in the fuzzy rule-based system, the crisp inputs are firstly converted to the fuzzy values by comparing the input crisp values with the database membership functions. Then IF-THEN fuzzy rules are applied on the input fuzzy values to make consequence of each rule, as the output fuzzy values. The outputs obtained for each rule are aggregated into a single output fuzzy value, using a fuzzy aggregation operator. Finally, defuzzification is utilized to convert the output fuzzy value to the real world value as the output.

Mamdani type is one the most commonly used fuzzy inference method which is employed in this study as well [11]. In this method, R_i as ith IF-THEN linguistic rules is defined by:

R_i: IF x_1 is A_{i1} and ... and x_r is A_{ir} THEN y is C_i, where x_j ($j=1, 2, \ldots, r$) are the input linguistic variables, y is the output linguistic variable, and A_{ij} and C_i are fuzzy sets for x_j and y respectively. Given input of the form: x_1 is A'_1, x_2 is A'_2, ..., x_r is A'_r, where $F'_1, F'_2, \ldots, F'_r$ are fuzzy subsets of $U_1, U_2, \ldots, U_r$.

The contribution of the rule R_i to Mamdani model's output is a fuzzy set whose membership function is:

$$\mu_{C'_i}(y) = (\alpha_{i1} \wedge \alpha_{i2} \wedge \cdots \wedge \alpha_{in}) \wedge \mu_{C_i}(y) \quad (1)$$

where α_i is the matching degree of the rule R_i and n is the number of input variables

The matching degree between x_j and R_i's condition on x_j, α_{ij}, is:

$$\alpha_{ij} = \sup(\mu_{A'_j}(x_j) \wedge \mu_{A_{ij}}(x_j)) \quad (2)$$

where $\wedge$ and sup denote the "min" and "supremum" operators, respectively.

The final output of the model is the aggregation of outputs from all rules using the "max" operator as:

$$\mu_C(y) = \max\{\mu_{C'_1}(y), \mu_{C'_2}(y), \ldots, \mu_{C'_L}(y)\} \quad (3)$$

where L is the number of MF's.

The output C is a fuzzy value, which cannot be applied directly. Therefore, it is necessary to convert the output fuzzy value into a crisp one which is achieved using defuzzification process.

2.2. Geometric invariant moment

Geometric invariant moment technique which is based on theory of the algebraic invariant, attempts to extract Rotation/Scale/Translation (RST)-invariant visual features. This type of features which was firstly introduced by Hu [12], has been successfully applied in aircraft identification, texture classification, radar images and optical images matching [13]. Two-dimensional moments of a digital image are given as:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad p, q = 0, 1, 2, \ldots \tag{4}$$

where $f(x, y)$ is the value of the digital image in (x, y) coordinate.

The corresponding central moment is defined as:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{5}$$

where:

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{6}$$

The normalized central moment of order $(p+q)$ is also defined as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{pq}^{\gamma}} \tag{7}$$

where:

$$\gamma = \frac{p+q}{2} + 1 \tag{8}$$

In particular, Hu [12] defines seven normalized central moments through the third order which are scale, position, and orientation invariant. In terms of the central moments, these seven moments are:

$$\begin{aligned}
\Phi_1 &= \eta_{20} + \eta_{02} \\
\Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\Phi_4 &= (\eta_{30} - \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\
\Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\Phi_6 &= (\eta_{20} - 3\eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{12} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} - \eta_{03}) \\
\Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{9}$$

3. The proposed system

The proposed system consists of four classifiers and a simple arbitration algorithm. The first classifier detects skin regions in the color input image, while the second classifier detects lips regions in the detected skin regions. The third classifier is designed to detect frontal/near frontal faces while the last one is designed to detect profile faces. Each classifier consists of an accurately designed FIS with proper input features and a likelihood output. Meanwhile, a simple threshold which is optimized by a genetic algorithm (GA) is applied to the likelihood output to finalize the detection. Finally, a simple arbitrary algorithm qualifies the outputs of the frontal/near frontal and profile classifiers to detect faces and non-faces regions as well as the face poses, including frontal/near frontal and profile ones.

3.1. Skin-color segmentation

HSI (Hue-Saturation-Intensity) color space as perceptual features of color is a very good choice for skin segmentation, since the transformation of RGB to HIS is highly invariant to high intensity at white lights, ambient light and surface orientations relative to the light source [3]. Using following non-linear equations, RGB color space transforms to HSI color space, where the intensity component is separated from the chrominance components:

$$\begin{aligned}
H &= \cos^{-1} \frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \\
S &= 1 - 3\frac{\min(R, G, B)}{R+G+B} \\
I &= \frac{1}{3}(R+G+B)
\end{aligned} \tag{10}$$

Although in many skin classification research the intensity component is omitted, 2D projection of 3D color distribution cannot completely adapts itself to lighting variation. Furthermore, simply discarding the luminance information affects the model's accuracy. To obtain optimum color space, different fractions of color space components are tried, and finally the color space $((1/2)H, S, (2/3)I)$ is chosen by trial and error. Then over more than one million pixels gathered from skin samples of different face database, the average of the chosen color space is computed as the skin vector mean.

After transforming the input image into the chosen color space, the next step is finding the skin pixels. Mamdani fuzzy system used, is a 1-input, 1-output system applying the normalized Euclidean distance between the color of each pixel to the skin vector mean as an input, and the likelihood of being skin pixel as an output. *Subtractive clustering* [14] is applied on input space (contain 132,000 skin and non-skin pixels) to decide on the number of membership functions (MF's) and rules. Utilizing the obtained four clusters information and experimental knowledge, input and output MF's are designed. The semantic meaning is assigned to each cluster for better understanding. The achieved rule in skin-color segmentation FIS is:

IF input is Z, THEN output is Z

where $Z \in$ {Skin, Rather Skin, Low Probability Skin, Non-Skin}.

To achieve the crisp output, centroid method is chosen, which is the most widely used one among all defuzzification approaches [10]. MF's of the input and output is depicted in Fig. 2 where the

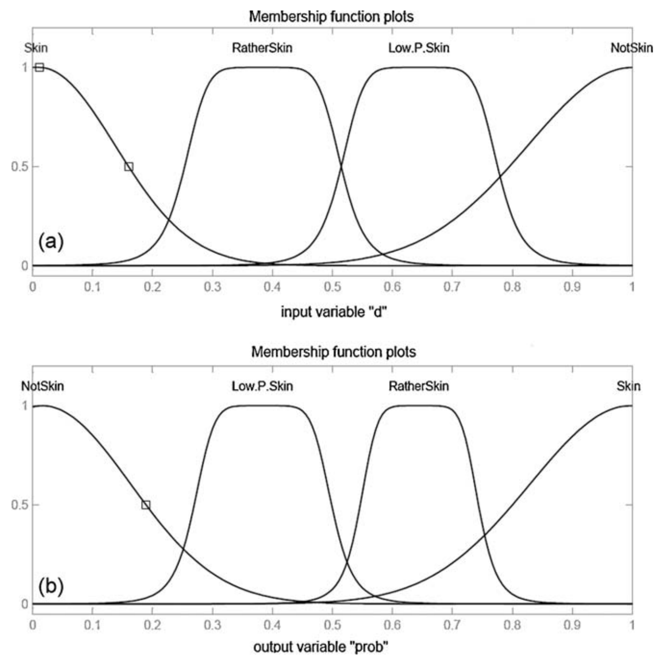

Fig. 2. (a) Input and (b) output MFs for the proposed skin-color segmentation FIS.

input "d" is the normalized Euclidean distance between the color vector of each pixel to the vector mean of skin-color space and the output "prob" is the skin-likelihood, between 0 and 1. The result of applying such a system is the skin-likelihood image; that is, the gray scale image whose gray values represent the likelihood of the pixel belonging to the skin.

To make a binary image, an appropriate threshold should be selected which is optimized by GA. GA is the most extended group of evolutionary technique known, which rely on the use of a selection, crossover and mutation operators [15]. The threshold is the chromosome of the GA, whose fitness function compares the whole detected skin pixels in the sample images with the actual number of these pixels, and attempts to minimize the different.

Over 45 randomly selected skin samples, the obtained threshold is 0.8964. It means that the pixel with skin-likelihood more than 0.8964 is regarded as the skin pixels. The binary image is formed by setting skin pixels to '1' and all other pixels to '0'. After that, morphological processing, consists of filling holes and opening followed by closing [16], is accomplished to acquire separate and connected regions. After labeling the connected components, each region is applied as an input for the next step.

3.2. Lips-color segmentation

The lips-color classifier searches the whole skin region to find lips pixels. In the proposed lips-color classifier, the color of pixels is transformed to the normalized RGB color space using Eq. (11):

$$r(x,y)=\frac{R(x,y)}{\mathrm{Max}(R,G,B)}$$
$$g(x,y)=\frac{G(x,y)}{\mathrm{Max}(R,G,B)} \qquad (11)$$
$$b(x,y)=\frac{B(x,y)}{\mathrm{Max}(R,G,B)}$$

This color space has provided the best results in finding lips [5].

Using the normalized RGB color space within the desired region, the normalized r–g values are utilized as fuzzy system input. The properties and design process of this system are similar with the skin-color segmentation algorithm. The input space (contain 50,700 lips and non-lips pixels) is divided into three clusters using the subtractive clustering.

The input r–g is recognized by linguistic terms {*Lips*, *Rather Lips*, *Non-Lips*} for both the input and the output, rules would be:

IF input is Z, THEN output is Z

where $Z \in$ {Lips, Rather Lips, Non-Lips}.

MF's of the input and output is depicted in Fig. 3 where the input "r–g" is the normalized r–g and the output "Lip.P" is the lips-likelihood, between 0 and 1.

Similar to the skin detection scheme, the result of this stage is lip-likelihood image. The computed threshold is 0.827 by GA over 42 randomly selected lips samples. The thresholding algorithm on

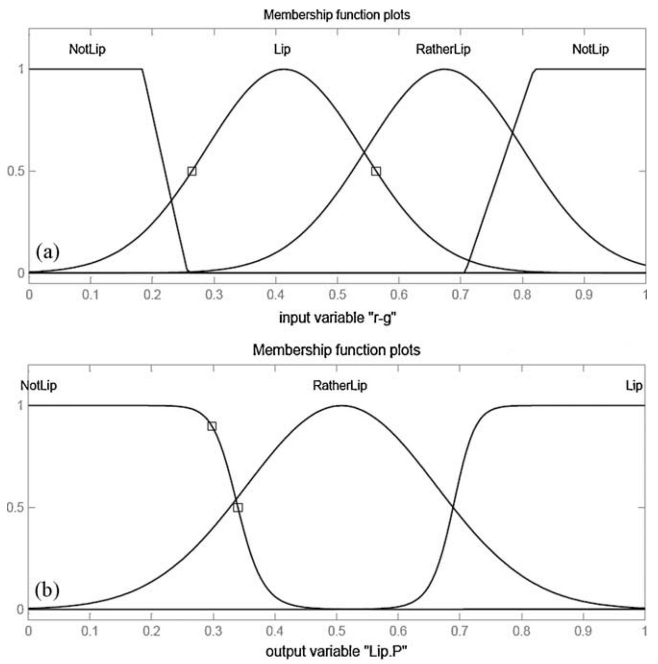

Fig. 3. (a) Input and (b) output MFs for the proposed lips-color segmentation FIS.

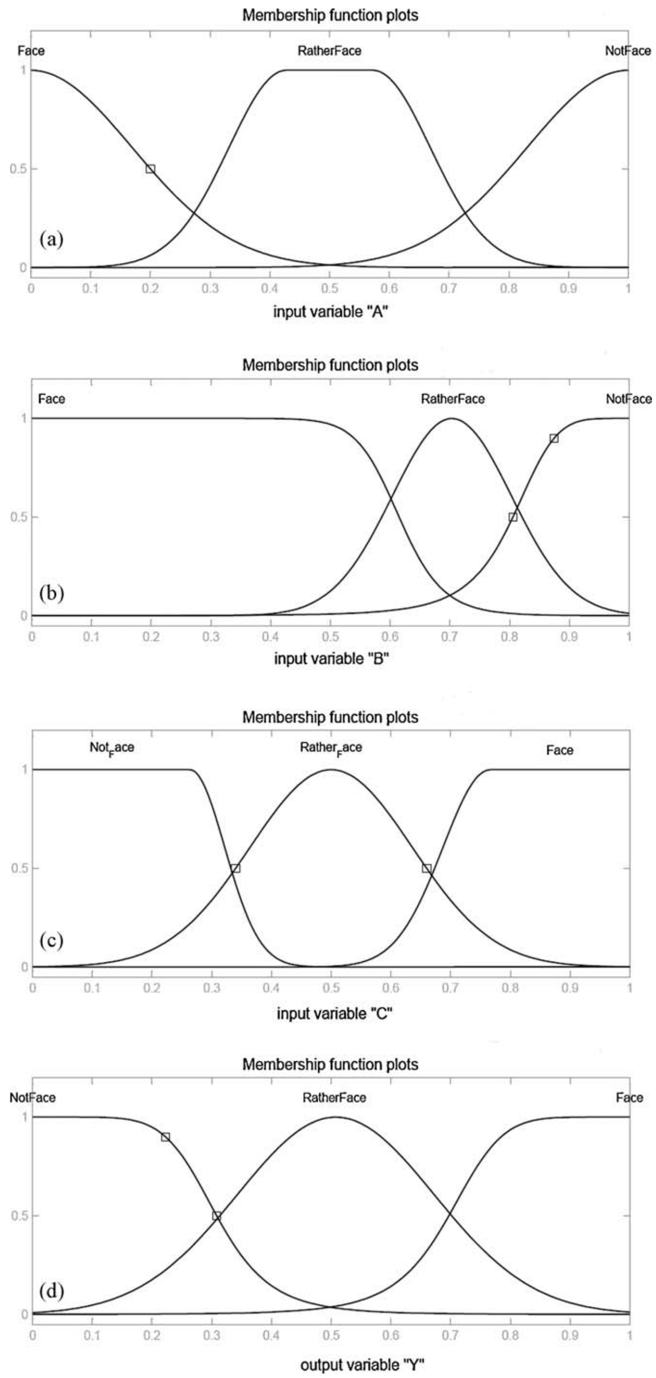

Fig. 4. (a–c) Inputs and (d) output MF's for the proposed frontal/near frontal face detection FIS.

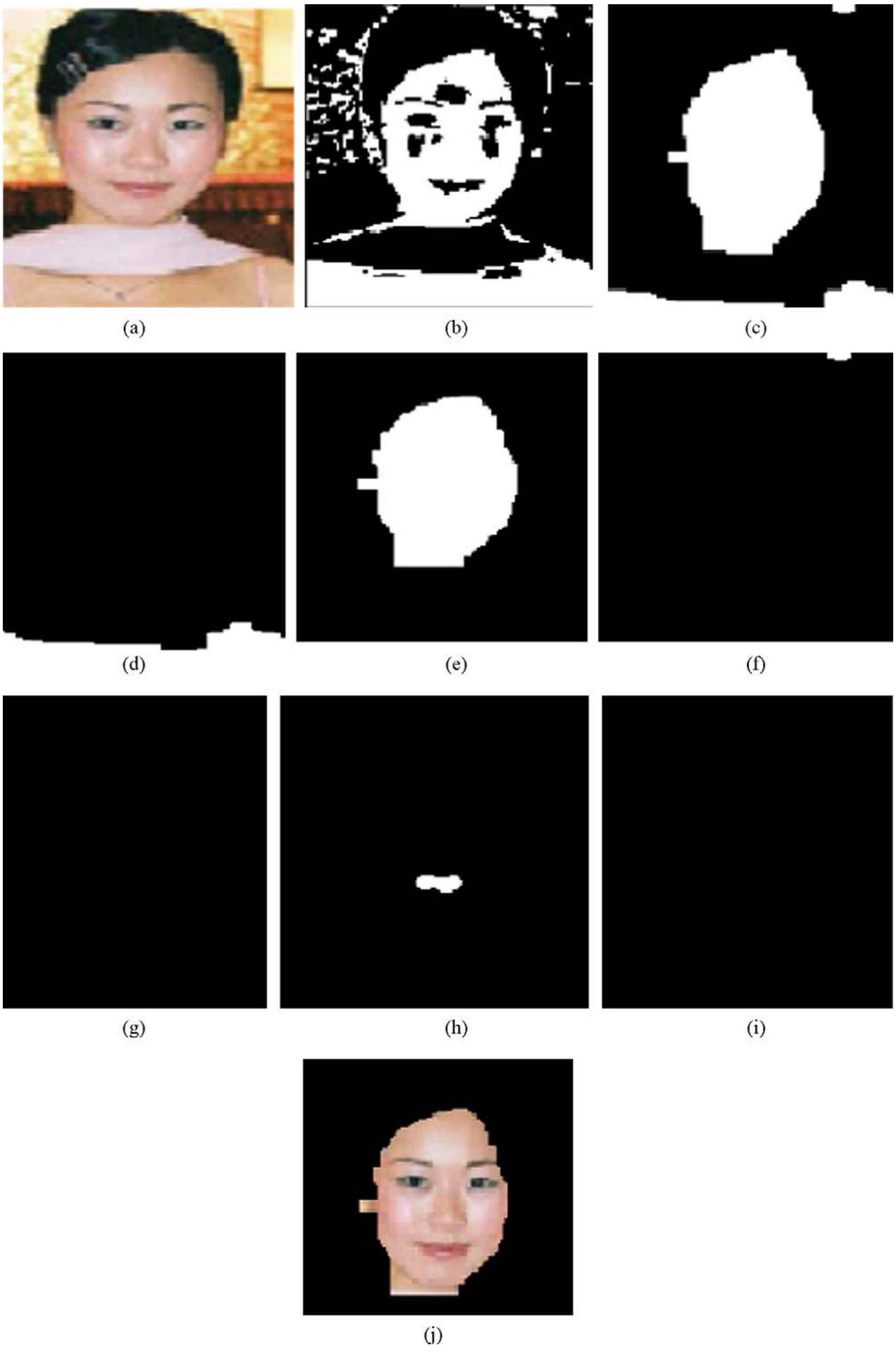

Fig. 5. (a) Main image, (b) skin regions, (c) skin regions after morphological processing, (d–f) 3 separated connected components, resulted after labeling skin regions, (g–i) output images of lips-color segmentation system, respectively for lips areas found in regions (d–f) in the normalized RGB color space, and (j) detected face area.

the lips-likelihood image makes binary image of the lips regions. Filling the holes and erosion plus opening by reconstruction [16], are the morphological processing performed at the end of this stage. There might be a large amount of r in skin-color, especially in cheek area. Applying lips-color segmentation algorithm just in the lower half skin region, is a new idea, to reduce the false positive in lips detection system.

3.3. *Frontal/near frontal face detecting algorithm*

In this algorithm, the frontal and near frontal face blobs would be selected in the skin regions. Decision-making system for frontal and near frontal face detection is a FIS with one output and three inputs. The first input "A" is the ratio of the lips area to the face area. The second input "B" is the tangent of angle between the maximum length line of face area and the minimum length line of the lips area. These two lines should be approximately parallel. Angles less than 30 are acceptable. The third input "C" reveals the similarity of region shape to an ellipse. This ellipse estimates the face shape. Applying such inputs, the false positive of algorithm would be the lowest possible value. Output "Y" expresses the likelihood of a region being face.

Utilizing 100 different face and non-face regions, three subspaces are introduced by the subtractive clustering algorithm to generate the input space. The inputs and output MFs are illustrated

Table 1
Geometric invariant moment (GM) values over 10 ear texture sample images.

GM	Sample images of ear									
	1	2	3	4	5	6	7	8	9	10
Φ_1	5.34	5.20	6.27	5.50	6.14	5.62	5.95	6.02	6.30	5.56
Φ_2	12.00	11.80	11.67	12.87	13.50	12.12	13.02	12.16	13.36	12.50
Φ_3	19.56	18.78	20.97	18.82	20.29	21.39	21.69	20.31	22.85	21.34
Φ_4	20.65	22.34	25.25	20.28	23.10	22.94	26.42	25.47	25.67	22.54
Φ_5	40.78	43.87	48.09	39.85	45.81	45.26	44.59	48.71	49.35	41.22
Φ_6	26.92	28.35	30.73	26.56	29.87	29.57	30.38	32.11	32.88	29.69
Φ_7	42.40	42.98	41.27	43.01	45.44	45.89	46.21	44.80	41.32	44.59

in Fig. 4. The linguistic terms are: {*Face*, *Rather Face*, *Non-Face*} for both input and output MFs. The defined rules are:

IF input is Z, THEN output is Z

where $Z \in$ {Face, Rather Face, Non-Face}.

Regions with more than 0.74 face likelihood are introduced as a frontal/near frontal face. The threshold is selected by GA over 42 randomly selected frontal and near frontal face images. Fig. 5 shows different steps of the proposed frontal/near frontal face detection algorithm applied on a sample image. This algorithm works successfully for frontal and near frontal faces, in which the lips area could be detected, but it has low detection rate for profile faces [17]. So, to compensate this disadvantage, another algorithm, a profile face detecting algorithm, is developed. More details are presented in the following sections.

3.4. Profile face detecting algorithm

In the proposed profile face detecting algorithm, the properties of ear texture are used. First of all, the ear location is estimated by its relative position to nose tip. The suggested process of finding nose tip in [18], is modified to be size invariant.

Suppose (x, y) is the corresponding coordinate of a pixel in a image. In the proposed algorithm for locating the nose, x values along each row which we first encounter a skin pixel in the binary skin image is searched. With the median of these values (x_{median}) the approximation x value of the face contour is obtained. Then, the median of y values in the range of x_{median} is calculated (y_{median}). In the last step, the x value of nose tip is introduced as the minimum of x values in the range of:

$$y_{\text{median}} - \frac{n}{2} < y < y_{\text{median}} + \frac{n}{2} \qquad (12)$$

where n is the number of y in the range of x_{median}. This method avoids the possibility of locating hair or chin as the nose tip.

After determining the x and y of the nose tip, the ear location is estimated. Since the y values of the noise tip and the ear are close, a horizontal search space which its y value is equal to y_{nose} is enough. Therefore, the x value of the search space is between the x value of the nose tip and the last skin pixel, in the same y value. On the other hand, the ear dimension is estimated using area of the detected face. In the search region of ear existence, the properties of texture are evaluated. Fig. 6 shows a sample image of the profile face, binary skin region, nose location and finally, the search region for estimating ear location.

After studying various texture classification methods, such as statistical properties of the intensity histogram (statistical moments), geometric invariant moment (GM) is chosen to evaluate the ear area. A set of 10 ear images is collected to investigate the moment values. As it can be seen in Table 1, Φ_1 and Φ_2 are sufficient and more suitable to verify ear texture. It should be noted that, these values reveal the absolute values of log of results, applying Eq. (9). Using log, dynamic range would be reduced and the absolute

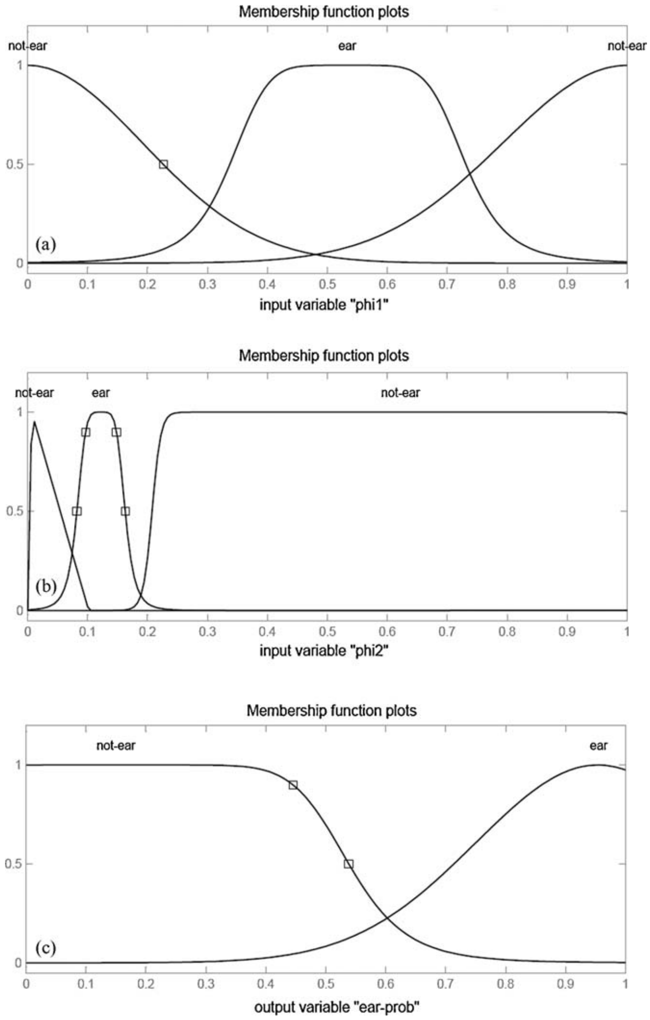

Fig. 6. (a) Input image, (b) skin region and nose tip position, (c) divide region to estimate ear position, and (d) detected face in the image by verifying ear texture.

values avoids having to deal with the complex numbers resulted when the log of negative moment is computed.

Again, instead of applying a crisp threshold, a FIS is designed based on the Φ_1 and Φ_2. These two values are employed as the inputs and the probability of ear texture existence is the output. Applying the subtractive clustering algorithm over 100 ear and non-ear image samples, results in two clear clusters, therefore {*Ear*, *Non-Ear*} are chosen as linguistic term where the rules in the proposed FIS would be:

IF input is Z, THEN output is Z

where $Z \in$ {Ear, Non-Ear}.

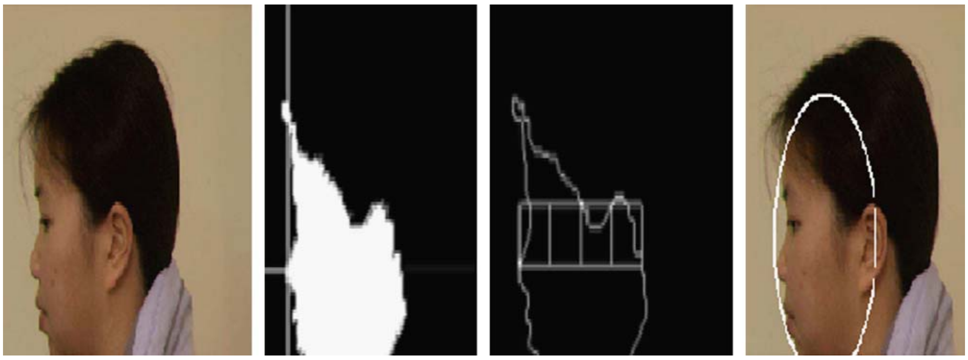

Fig. 7. (a and b) Inputs and (c) output MFs for the proposed ear texture detection FIS.

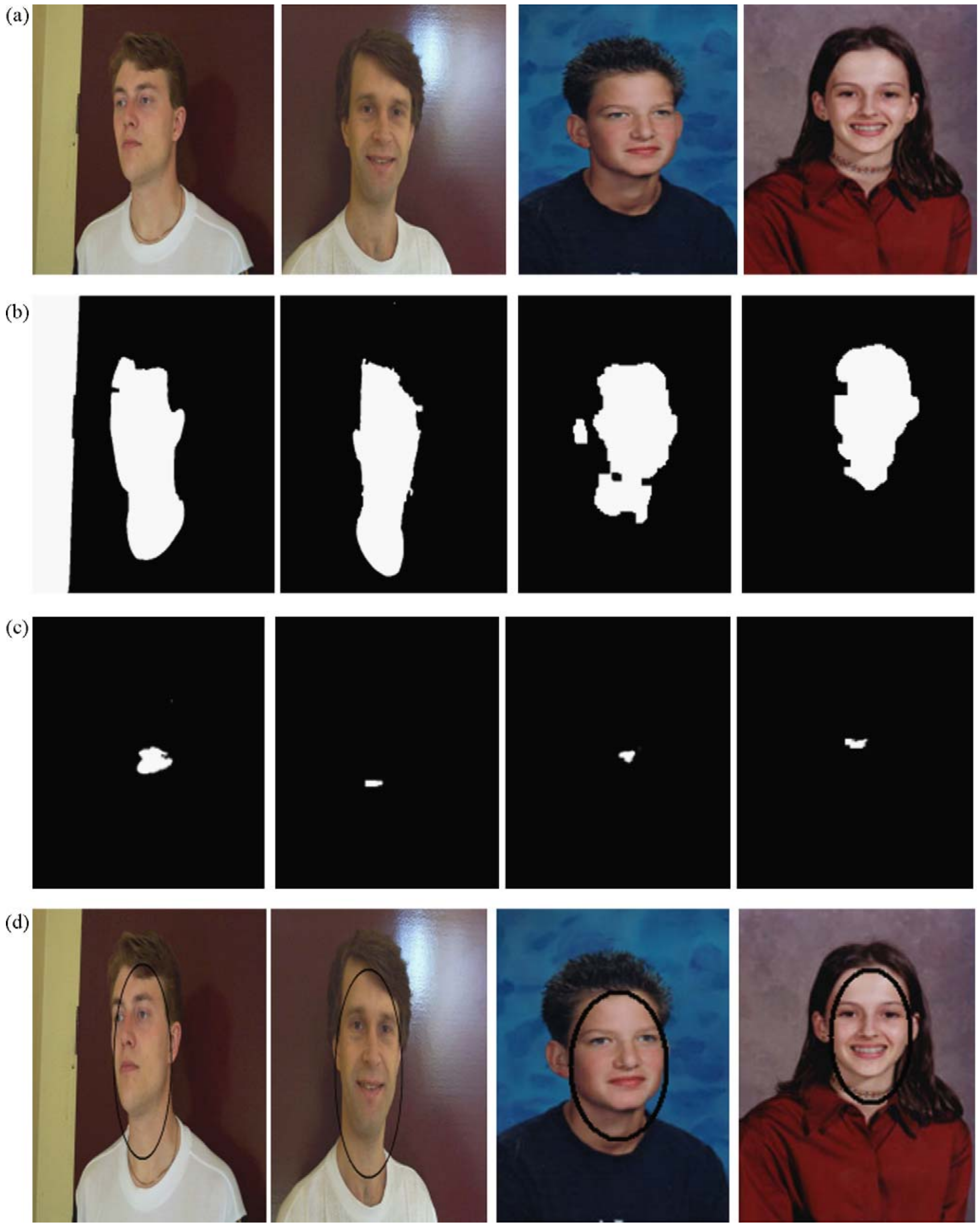

Fig. 8. Examples of applying the proposed face detection algorithm on some images of the first experiment (a) main images, (b) skin regions, (c) lips regions, and (d) face detected areas.

Table 2
Experimental results on HHI and Champion databases (DR: detection rate).

	Face database					
	HHI database					Champion database
	Head poses				Total	Total
Algorithm	Frontal 66[a]	Near frontal 54[a]	Half profile 75[a]	Profile 11[a]	206[a]	227[a]
The proposed method (DR%)	93.9	96.3	94.6	90.9	94.7	97.4
Jing et al. (DR%)	92.4	98.1	92.0	90.9	93.7	95.2
Hsu et al. (DR%)	88.4	90.7	74.7	18.2	80.6	91.6
Mousavi et al. (DR%)	87.9	88.8	65.3	18.2	76.2	92.5

[a] No. of image.

Table 3
Experimental results over the different face databases (DR: detection rate).

Dataset	Head pose		
	Frontal	Near frontal	Profile
Bao [21]			
Number of images	40	20	5
DR (%)	95.0	95.0	80.0
Caltech [22]			
Number of images	80	–	–
DR (%)	98.8	–	–
Essex [23]			
Number of images	70	10	–
DR (%)	100.0	100.0	–
IMM [24]			
Number of images	60	40	–
DR (%)	100.0	97.5	–
Profile images [18]			
Number of images	–	–	25
DR (%)	–	–	92.0

Fig. 7 shows the MF's of the proposed FIS where "phi1" and "phi2" are the normalized inputs geometric moment Φ_1 and Φ_2 and "ear-prob" is the output variable. The optimized threshold 0.67 is selected by GA over 36 randomly selected ear images. If the existence of ear in the region of interest is verified (ear-likelihood is more than 0.67), the blob is known as the profile face. Fig. 7 depicts the described stages.

3.5. Final arbitration algorithm

In final arbitration stage, both profile and frontal face detecting system are used to detect faces. So, each skin blob is entered to both systems. Depending on the systems response, three different conditions may occur:

1. If none of the frontal or profile systems identify the skin region as a face, it would be discarded.
2. If one of the systems verifies the skin region as a face, that blob would introduce as the frontal or profile face.
3. If the outputs of both systems are positive for one skin region, it is a face but corresponding pose is unknown.

The regions lie within the third group, could be announced as face, but to increase the accuracy of the proposed system, a pose detection approach is designed and applied for such cases. The position of lips area in face region is estimated. The center of x values of face and lips areas is calculated and their distance is called d_1. Meanwhile, d_2 is defined as the length of face region minor axis. Regarding these two values, if $d_1 < d_2/4$ the detected face is frontal; otherwise, it is a profile one.

4. Experimental results

To evaluate the performance of the proposed face detection system, some experiments are carried out with different face databases. In the first experiment, the HHI face database [19] including 206 images and Champions face database [20] including 227 images, are used. Table 2 shows the detection rates for the proposed method along by the results of Hsu et al. [7], Jing et al. [9] and Mousavi et al. [17] face detection algorithms. The last compared algorithm [17] is another fuzzy rule-based system which is used FIS's to segment skin-color and then detect the faces based on some extracted features. As there are the images with one face and simple background in these databases, a remarkable result achieved. It should be mentioned that the images lie within the half profile group, mostly detected by frontal/near frontal detection algorithm and some of them (if the nose tip was detectable) detected by the profile face detection algorithm. Fig. 8 illustrates some output examples of the proposed scheme.

To demonstrate further effectiveness of proposed approach, a dataset is formed, containing 350 face images. Various pose/lighting condition/facial expression/size (125×93–296×448 pixels), and images of people with different ethnic origins are included. These images are gathered from:

1. Bao image database includes 370 face images from various races, mostly from Asia, with wide range of size and pose [21].
2. Caltech frontal face database, collected by Markus Weber at California institute of technology. There are 450 face images from 27 persons taken under different lighting/expression/backgrounds [22].

Fig. 9. Examples of applying the proposed face detection algorithm on some images of the second experiment with various pose and background.

3. Essex University image database contains images of 395 individuals, classified in four groups by their variation of background and scale, versus extreme variation of expressions [23].
4. The IMM face database contains 240 still images of 40 different human faces, including 7 females and 33 males with various facial expressions and poses [24].
5. Due to the lack of profile face images in above databases, 25 profile faces images are added from [18].

In the second experiment, the performance of the proposed method is investigated among all the images which are divided into three groups, according to their frontal, near frontal and profile pose. Table 3 shows detection rates for different datasets and poses. 98.8% is an acceptable detection rate, obtained for Caltech database, which contain just frontal face images. For Bao database, the detection rate is slightly lower. Considering wide range of skin and lips-color of people of different origins and face images with various sizes included in this dataset, the obtained results are still promising. For images from Essex University and IMM databases, 100 and 97.5% detection rates are achieved respectively for frontal and near frontal faces. So, it could be concluded that the proposed method has approximately no dependency on facial expressions.

Table 4
Summarized experimental results over all 783 face image samples (DR: detection rate).

Head pose	Frontal	Near frontal	Profile
Number of images	543	199	41
DR (%)	97.6	96.0	90.2

In Table 4, the average results for all images including 783 face images are summarized. 97.6, 96.0 and, 90.2% correct detection rates, for frontal, near frontal and profile faces show that the proposed method is the precise and robust face detector in color image. Fig. 9 illustrates some more results obtained from MATLAB simulation as well.

5. Conclusion

In this paper, a novel fuzzy rule-based method to detect human faces in color images was proposed and described, where two accurately tuned FIS's were employed respectively to create skin-likelihood and lip-likelihood image. Then two other accurately tuned FIS's were employed to detect frontal/near frontal and profile face blobs among the skin regions. The number of rules was selected by the subtractive clustering and the thresholds were determined by a GA.

Along with the color information, the position and relative area of lips, and approximated face shape for frontal faces, and ear texture properties for profile ones, were used to verify the detected faces.

The obtained results were presented for HHI and Champion face databases and compared to three face detection algorithms. In addition, the proposed system was applied on wide range of face images from different databases and discussed the achieved detection rates were presented and discussed. The experimental results showed that the proposed system works well in detection of various face in different poses.

As mentioned before, detecting the profile, near frontal and frontal faces using a single algorithm, is an important advantage of the developed system. Moreover; this algorithm, could be used as the pose detection scheme as well. In final stage of a decision-making process, it could be easily decided that the detected face is either frontal or profile. We are going to estimate the face turn angle in images with a little change in the future.

To devise a system that can adapt itself with more view changes, adding more features such as shape and intensity information to the lips detection system, and using other features of the profile face shape to improve the detection rate are possible ways to try.

References

[1] M. Yang, D.J. Kriegman, N. Ahuja, Detecting face in image: a survey, IEEE Transaction on Pattern Analysis and Machine Intelligence 24 (1) (2002).

[2] H. Rowley, S. Baluja, T. Kanade, Neural network-based facedetection, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1996, pp. 203–208.

[3] P. Kakumanu, S. Makrogiannis, N. Bourbaki, A survey of skin-color modeling and detection methods, Pattern Recognition 40 (3) (2007) 1106–1122.

[4] N. Eveno, A. Caplier, P.Y. Coulon, A new transformation for lips segmentation, multimedia signal processing, in: 2001 IEEE Fourth Workshop, 2001, p. 3-8.

[5] J.A. Dargham, A. Chekima, Lips Detection in the Normalised RGB Color Scheme, Information and Communication Technologies, ICTTA06, vol. 1, 2006, pp. 1546–1551.

[6] J.-C. Terrillon, M. David, S. Akamatsu, Detection of human faces in complex scene image by use of a skin color model and of invariant Fourier-Mellin moments, pattern recognition, in: Proceeding Fourteenth International Conference, vol. 2, 1998, pp. 1350–1355.

[7] R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color image, IEEE Transaction on Pattern Analysis and Machine Intelligence 24 (2002) 696–706.

[8] V. Vezhnevets, V. Sazonov, A. Andreeva, A survey on pixel-based skin color detection techniques, in: Proc. Graphicon, 2003, pp. 85–92.

[9] M.-Q. Jing, L.-H. Chen, Novel face-detection method under various environments, Optical Engineering 48 (6) (2009).

[10] J. Yen, Fuzzy logic–a modern perspective, IEEE Transactions on Knowledge and Data Engineering 11 (1) (1989) 153–165.

[11] S.N. Sivanandum, S. Sumathi, S.N. Deepa, Introduction to Fuzzy Logic Using MATLAB, Springer-Verlag, Berlin/Heidelberg, 2007.

[12] M.K. Hu, Visual pattern recognition by moment invariants, IRE Transactions on Information Theory 8 (1962) 179–187.

[13] M. Rizon, H. Yazid, P. Saad, A. Yeon Md Shakaff, A. Rahman Saad, Object detection using geometric invariant moment, American Journal of Applied Sciences 2 (6) (2006) 1876–1878.

[14] A. Priyono, M. Ridwan, Generation of fuzzy rules with subtractive clustering, Jurnal Teknologi 43 (D) (2005) 143–153.

[15] S.N. Sivanandam, S.N. Deepa, Introduction to Genetic Algorithms, Springer-Verlag, Berlin/Heidelberg, 2008, pp. 165–209.

[16] R. Davies, Machine Vision, Morgan Kaufman, 2004.

[17] S. Mousavi, P. Moallem, M. Kavosh, Detecting faces in colour images using fuzzy rule based classifier, in: 2nd Joint Congress on Fuzzy and Intelligent Systems, Tehran, Iran, 2008.

[18] P. Yan, K.W. Boyer, Biometric recognition using 3D ear shape, IEEE Transaction on Pattern Analysis and Machine Intelligence 29 (8) (2007).

[19] MPEG7 Content Set from Heinrich Hertz Institute. http://www.darmstadt.gmd.de/mobile/hm/projects/MPEG7/Documents/N2466.html, 1998.

[20] The Campion Database. http://www.libfind.unl.edu/alumni/events/breakfastt_for_champions.htm, 2001.

[21] http://www.facedetection.com.

[22] http://www.vision.caltech.edu.

[23] http://cswww.essex.ac.uk/mv.

[24] http://www.imm.dtu.dk/~aam/aamexplorer.