ER7ST-Algorithm for Extracting Facial Expressions

Ahmad Tayyar¹, Shadi Al-Shehabi², and Majida AlBakoor³ ¹Department of Computer Science, Jerash University, Jordan ²Department of C omputer Engineeringm, Türk Hava Kurumu Üniversitesi, Turkey ³Department of Mathematics, Aleppo University, Syria

Abstract: This paper, proposes a new algorithm for recognition of facial expressions, called ER7ST. Studied expressions are anger, disgust, fear, happiness natural, sadness and surprise. Proposed method is based on extraction of the essential objects, then finding of characteristic points positions of each object. Detected points slant and its average are calculated on the basis of fixed points. Mug database is considered as data source for training and testing. We collect set of images. ER7ST is designed to define the work area of face that contains characteristic expressions, which its centre is face centre and its dimensions are 8×16 supposed squares. ER7ST algorithm discovers the essential objects depending on the coordinate of formed ellipse on studied object. Object gradient is ranked between [-60, +60] degrees. Our algorithm detects Convex Hull points upon detected objects and then its slant is calculated. Slant vectors are formed; some calculations are done to be a good input to network. Net contains input, three hidden layers and output layer. After training on set of faces and testing on new data, the recognition rate is ranked between 60% and 95%, experiments have shown that method was efficient and results were very encouraging in this field, especially network can be trained on new situations of expressions.

Keywords: Facial expressions, feature extraction, Essential objects, slant.

Received September 24, 2014; accepted February 5, 2015; published online July 11, 2015

1. Introduction

Face recognition is one of the most intensively searched topics in biometrics, due to many applications in the field of our daily life. For that, many of facial recognition methods appeared which they differ according to the manner of representation, classification and discrimination [4, 7, 15]. However, the attention of the scientific research has gone widely towards focusing the efforts on the topic of emotions and nervous situations of human beings and thus the study of facial expressions to design an effective and capable system for recognition of facial expressions [10, 14, 21]. In this area, many methods appeared and efforts have been made to enrich this field. Most of the research was based on the use of fixed images or isolated images from videos captured through their experiences [13, 17]. Many researchers focus in this field on invention new techniques to recognize the facial expressions being the most important item in this scope; This in turn will lead to discernment the human psychological situation. Some researchers are using Gabor filters to extract facial expressions to erect input of Neural Network (NN) [3, 9]. As others tend to design 3D model from 2D face image to extract information from its intention to recognize emotions and psychological situation of human [1, 16]. In addition, the others tend to use models as GLTP model for collection data from encryption of local positions from mini models of GLTP, which they compose beam represents facial expressions [12, 22]. Others have focused on the recognition issue, they compared

between the effectiveness of the inferential system ANFIS and the effectiveness of back propagation NNs in the possibility of discrimination facial expressions [11, 20]. In our paper, we have designed algorithm that extracts essential objects of the face, which represents the facial expressions and then it defines characteristic points Convex Hull¹, it calculates slants of important points and its rates on the basis of two fixed points. First point is the centre of the Essential object; the second point is the centre of the face. Then, many complex calculations are performed to get a good input of NN. The recognition rate of the back propagation NN (for recognition of different facial expressions) in our method is a high and with lowest cost of time. A short comparison in this paper between our approach and similar approaches is presented in Table 1, with additional remarks below.

Table 1. Comparison with other approaches.

Database	Approach	Features No.	Expr.	Input	Reco. Rate	
Cohn-Kanade [10]	Gradient Local Ternary (GLTP) Pattern		6 (anger, disgust, fear, happiness sadness and surprise)	480×640, grayscale 1224 images	97.2 %	
Yale [12]	3D Model andgeodesic distance		5 (fear, happiness natural, sadness and surprise)	grayscale	Of 85% To 90%	
JAFFE [14]	ANFIS system and BPNN		7 (anger, disgust, fear, happiness natural, sadness and surprise)	grayscale 213 images	95.29 %	
MUG (ER7ST)	Extract Essential objects and Slant track of characteristic points		7 (anger, disgust, fear, happiness natural, sadness and surprise)	Color, grayscale 2912 images	Of 60 % To 95 %	

¹ set of points which form a convex polygon

ER7ST-Algorithm for Extracting Facial Expressions

In our paper, face image is taken and then set of processes is done to improve this image. The first step, convert the image from colour image to grayscale (if the image is colour). Density is adjusted based on the minimum and maximum values of it. Thus, enhancing of the contrast using histogram equalization. The second step is filtering the image depending on Median filter, which helps to remove the noise and to enhance the image edges. The image is rotated around its centre with suitable angle according to the circumstances of image capture. Next step is to discover the edges of studied image using the edge detection canny. After detecting the edges, we are working to discover the essential objects according to the proposed algorithm; we called it (ER7ST). The algorithm defines the region (work area) of expressions and its centre. Area is represented by a grid that its dimensions are 16×8 supposed squares; its centre is face centre. Algorithm discovers Essential objects which represent facial expressions, each object is characterized by its location within that area and the length of perimeter of this object is greater than the length of supposed pixels grid square. The algorithm discovers the Convex Hull points on studied object considering the direction is counter-clockwise through searching about these important points.

The slants of detected points and their averages are calculated on based the centre of studied object, the same is done on based the centre of face. Set of complex calculations is computed on these resulted data and is scaled using certain way to form a good input of neuron network. The results are 124 features that represent facial expressions and they are a good input of NN. The NN contains three layers; the input layer is composed of 124 neurons, three hidden layers and output layer. Last layer is composed of seven neurons which represents seven situations, they are anger, disgust, fear, happiness, natural, sad and surprise state. The net was trained on set of input data. Input data images were taken from mug database. After training, the net it tested on new set of images. The rate of recognition is ranked between 60% and 95%.

2. Pre-Processing Steps of Face Image

Face image captured by digital camera or any other tool, the image requires advanced pre-processing steps to improve the image quality [5]. The first step is to convert from colour image to grayscale image (if it is colour). The second step is image normalization by adjusting the intensity using (imadjust) function [6] then enhancing the contrast with histogram equalization. The third step is filtering using median filter in order to reduce noise (to be the image a good goal) [19] as shown in Figures 1-a and d. In the fourth step, the image is rotated (if it requires) by a suitable angle around it, the angle is defined according to style of photography. The fifth step detects the edges by canny method considering the threshold is 0.08 and sigma is 1 as shown in Figures 1-e and f.

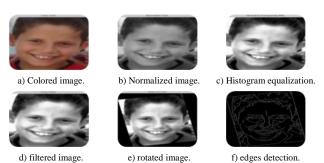


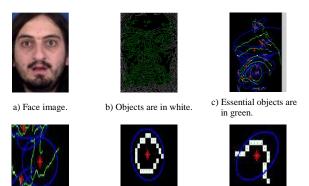
Figure 1. Face image pre-processing.

3. Face Image Processing

Set of steps are done to improve the face image, then the image processing is next stage to extract the facial expressions. The face shape is extracted according to a new proposed algorithm; we called it "ER7ST²".

We consider that the horizontal lines (edges) are more influential to discriminate the nervous situations and to understand the facial expressions, so the idea is to find this type of edges, after the edges are detected.

The work area and its centre are determined and represented as 16×8 supposed squares. Characteristic objects are discovered. Studied object is enclosed in ellipse. The Essential objects are extracted using the minimum and maximum coordinates of each object and the gradient between the x-axis and major axis is between -60 and +60 degrees (we supposed these values by experimentation) as shown in Figure 2-b. The characteristic objects boundaries are drawn in blue, some of characteristic objects are labelled as essential objects.



d) Object boundaries are in e) Object centre are in f) Formed ellipse on studied green. red. object are in blue.

Figure 2. Discovering of objects and formed ellipse on studied object.

3.1. Steps of Proposed Algorithm (ER7ST)

Essential objects are detected using slant transformation. The algorithm selects only important horizontally elongated objects, therefore significantly reduces the computational complexity. The algorithm steps can be summarized in two stages: The first stage is extracting of characteristic objects and the second stage is extracting essential objects.

²An abbreviation for the phrase following Expressions Recognition 7 Slant Tracking.

3.1.1. Stage of Extracting the Characteristic Objects

Characteristic objects are detected depending on the following:

- 1. Centre of face image $M(X_k, Y_k)$ is found.
- 2. Work area is calculated and its centre is detected. The area is defined as the middle of face space containing the eyes, nose and mouth. That space is represented by about three-quarters of the face. This area is divided into 16×8 supposed squares. The square length is denoted as Square_line_Y.
- 3. Exterior boundaries of objects are traced, with any holes inside these objects neglected. These objects are parents while their children boundaries represent continuous regions in the studied image.
- 4. For each object and its boundaries, the analysis is performed as follows:
 - The centre of object is located and denoted as Center_Poi.
 - The minima of coordinates on X-axis and Y-axis are calculated, they are denoted respectively as X-obj-min and Y-obj-min. Similarly, the maxima of coordinates are denoted respectively as X-obj-max and Y-obj-max.
 - The horizontal line horz_line is represented by the difference of maximum and minimum in Y_axis.
 - The vertical line_vert line is represented by the difference of maximum and minimum in X_axis.
 - An ellipse on the studied object is formed, then major and minor axis lengths of ellipse are calculated. They are denoted as Minor Axis Length and Major Axis Length. Figures 2-c, d, e and f shows the ellipse on studied object.
 - The slant between the x-axis and the major axis of the ellipse is calculated, it is called obj_slant.
- 5. If the horz_line or the Major Axis Length is larger than the vert_line then the object is a characteristic object.

Characteristic objects can be reduced to essential objects that represent the most important information, which could enable us to realize the human situation. Moreover, Figure 3 displays characteristic and essential objects.







b) Characteristic objects boundaries are in blue and yellow.

aries c) Object centre is in red essential objects boundaries are in green.

Figure 3. Characteristic and essential objects.

- 6. Characteristic object is promoted to an essential object if at least one of the following conditions applies:
 - The horz_line is larger than the Square line_Y.
 - The Obj_Slant is larger than -60 and the Obj Slant is less than +60.
 - The horz_line or MajorAxisLength is larger than the line_vert.
- 7. Slant track is studied for classification of Essential Objects. Slant is calculated between ConvexHull points and two fixed points (first the Essential object centre and the second point is face centre).
 - For each detected essential object, the following steps are performed:
 - a. The Convex Hull points are found. They are denoted by ConHull_poi.
 - b. The search for start and end points of Essential object starts from the first quarter and goes counter clock wise, after the face image is divided into four-quarters. The start point is considered if the sum of eight neighbouring pixels is equal to 1. It is denoted by Start_Poi, the end point is denoted by End_Poi. The Figure 4 shows these points.

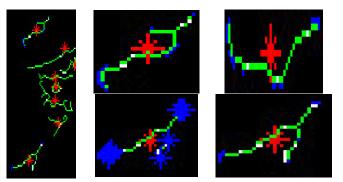


Figure 4. Examples of essential objects and Convex Hull points. (essential objects boundaries are in green and its centre is in red. Convex Hull points are positioned on essential objects in blue.)

• The slant between ConHull_poi point and the centre of face is calculated as in Equation 1, and the Figure 5-a explains that. as well as the way to perform required calculations.

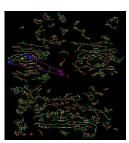
$$Slant 1 = \frac{(ConHull _Poi _Y - Center _i _Y)}{(ConHull _Poi _X - Center _i _X)}$$
(1)

• The slant between ConHull_poi point and the centre of essential object as in Equation 2 and Figures 5-b and c explains that.

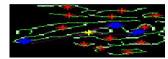
$$Slant 2 = \frac{(ConHull _Poi _Y - Center _Poi _Y)}{(ConHull _Poi _X - Center _Poi _X)}$$
(2)

• The slant average is calculated as in Equation 3:

$$MSlant = (Slant 1 + Slant 2)/2$$
(3)



a) Convex Hull points are in blue, object and face centre is in red.



b) Arrows are in purple colour refer to distance between Convex Hull point and essential object centre.



c) Arrow in in red colour refer to movement direction.

Figure 5. Distance between Convex Hull points and essential objects centres (face centre).

- The steps from b to d are repeated for all Convex Hull points of objects.
- The steps from a to f are repeated for every essential object. The result is a slants vector that is scaled into 124 features, to be a good input of NN and the following flow chart explains the proposed algorithm.

The images are taken from Mug Database (plus a set of images prepared by us) and are applied for seven expressions (anger, disgust, fear, happiness, natural, sadness and surprise). Then, the characteristic points of essential objects are discovered as 124 features and are processed by NN.

3.1.2. Data Extraction Represent Face

The resulted slant and its rates are represented by 124 features some calculations are done on it to get a good input for NN. The resulted data are scaled to get a good input of back propagation NN and Figure 6 shows the data represent the corresponding face image and the data after scaling.

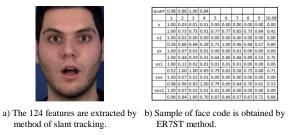


Figure 6. Data represents facial expressions according to ER7ST.

4. Configuration of Back Propagation Artificial Neural Networks

In our method, the Convex Hull points are detected, the slants are calculated between defined points and

the centre of Essential objects, then the slants are computed between that points and face centre. In addition, the averages of slants are calculated too. Data are scaled to 124 features that represent a good input data of back propagation NN [8]. Those values (features) are ranked between [0-1].

The NN contains the input and output layers, and three hidden layers. The input layer consists of 124 neurons. The neurons numbers of three hidden layers are set up according to these values in each layer sequentially: 550; 500; 500 neurons, these values were adopted through experimentation. The neurons number of output layer is seven neurons which represent various face expressions anger, disgust, fear, happiness, natural, sadness and surprise. The training parameters of the NN are set up according to these values: The momentum coefficient is 0.25, the learning rate is 0.05 and the sum square error is 0.00000585. This NN uses bipolar function as an activation function. The NN weights and biases are generated randomly. The values range of the input vector is between -1 and +1. Table 2 shows the proposed output of NN.

Table 2. Proposed output of NN.

Output of NN	Expression
[100000]	Anger
$[0\ 1\ 0\ 0\ 0\ 0]$	Disgust
[001000]	Free
[0 0 0 1 0 0 0]	Happiness
[0000100]	Normal
[0 0 0 0 0 1 0]	Sadness
[0000001]	surprise

The training and testing images are taken from Mug database (Mug is a database which contains images of 52 persons as samples. Seven of expressions are studied and eight images are taken for each expression for everyone (4 for training + 4 for testing)). The studied expressions are anger, disgust, fear, happiness, natural, sadness and surprise. Mug was collected under supervision of professor Anastasios N. Delopoulos of Aristotle University of Thessaloniki [2]. In addition to, the children images set picked up by us in one of the preschool in variety cases. These images play important role in training and testing of back propagation NN for recognition of several facial expressions. This type of networks is depending on training with method of teaching under Supervision using set of training pairs. Every pair consists of input and output vectors. These pairs introduced to the network. One of them provided to the network in each period that will produces output vector, then the difference is computed between the square of output and target vectors. After that, the error will return back to previous layer to adjust the weights on the base of the computed error, and so the training process was repeating until the square of difference between the output and target vectors reach to a small value which has been identified previously [18] and Figure 7 shows NN training.

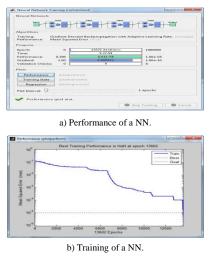


Figure 7. The parameters and training of ANN.

5. Results

The network is trained on the 1456 images, the number of Cycles is 13602 and the recognition rate is 100%. Its performance is tested by passing new data, which its number (1481). The rate of recognition is ranked between 60% and 95%.

Table 3 shows the results of NN testing on 28 new samples that can be split into seven groups; the first group of anger faces is represented by samples $(1 \dots 3)$; the second group of disgusted faces is represented by samples $(4 \dots 6)$; the third group of fear faces by samples $(7 \dots 9)$; the forth group of happy faces by samples $(10 \dots 12)$; the fifth group of natural faces by samples $(16 \dots 15)$; the sixth group of sad faces by samples $(16 \dots 18)$, the seventh group of surprised faces by samples $(19 \dots 21)$. Results of correct recognition are coloured in grey.

We notice through our experiments, that the network makes mistake in recognition between three cases (anger, sadness and fear). So that, the recognition rate was low sometimes for some of samples in anger situation as it reached to 60%, On the other hand, the recognition rate was the highest for the samples in happiness situation as it reached to 95%. The used network in our paper can be developed to lead the improvement of the discrimination by adding new samples for training with the possibility of training for new types of facial expressions and with high rate of recognition.

Table 3. The recognition results of testing the NN on set of new data.

	Ang	Ang	Ang	Disg	Disg	Disg	Fear	Fear	Fear	Happ	Happ	Happ	Neut	Neut	Neut	Sad	Sad	Sad	Surp	Surp	Surp
Anger	0.99	0.89	0.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00
Disgust	0.00	0.00	0.00	0.99	0.97	0.55	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fear	0.00	0.00	0.00	0.00	0.00	0.00	0.97	1.00	0.61	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.97	0.34	0.00
Happiness	0.00	0.00	0.24	0.00	0.58	0.00	0.10	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00
Neutral	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.88	0.83	0.98	0.22	0.00	0.00	0.00	0.00	0.02
Sadness	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.87	0.00	0.00	1.00	0.94	1.00	0.00	0.00	0.00
Surprise	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.92	0.78	1.00

6. Conclusions

In this paper, we proposed a new method to extract and classify facial expressions. We called it ER7ST. This algorithm is based on the election of the horizontal lines within the range [-60, +60] and consider it is the most important objects in the face. In the beginning, we have conversion the image to grayscale in case it is colour image, and then adjust the density and contrast to improve the image and then we worked on the purification of the image noise using the median filter. Then, we rotate the image according to the angle of image capture, and then we worked on discovery of the edges using the edge finder Canny.

Then, algorithm ER7ST starts to set the work area of the face, which includes the eyes and mouth, its centre is centre of the face; its dimensions are 8×16 and then discovery of essential objects in it depending on the minimum and maximum coordinates of points for each object. The algorithm searches for essential objects that its perimeter length of each object is greater than the length of the major- axis of formed ellipse on studied object. Then, Convex Hull points on these discovered objects is defined, then slant is calculated between these points and the centre of the characteristic object, and then between these points and the centre of the face and finally the slants averages are calculated. The algorithm performs some complex calculations on these slants vectors to be converted to the suitable input of back propagation NN. The network consists of input layer, three hidden layers, and output layer, which represents seven situations of human. They are anger, disgust, fear, happiness, natural, sadness and surprise. The network is trained on a set of faces and then is tested on a new set of images. The rate of recognition was good. The Mug database is considered as source of training and testing data. Thus, we can say that the process of extracting the facial expressions has taken an important place in this research. The idea of extracting the important and essential objects is done and then points Convex Hull are detected and then the slants are calculated between them and the fixed points. The slants vectors are converted to a good input of NN. Our algorithm is distinguished by a high rate of recognition with a speed of implementation and quality performance. Moreover, we can work on the development of research in this orientation and reach to recognize new patterns of facial expressions.

References

- [1] Ahmed F. and Hossain E., "Automated Facial Expression Recognition using Gradient-Based Ternary Texture Patterns," *Chinese Journal of Engineering*, vol. 2013, pp. 1-8, 2013.
- [2] Aifanti N., Papachristou C., and Delopoulos A., "The MUG Facial Expression Database," in Proceedings of the 11th International Workshop on Image Analysis for Multimedia Interactive Services, Desenzano, Italy, pp. 1-4, 2010.
- [3] Bhadu A., Tokas R., and Kumar V., "Facial Expression Recognition Using DCT, Gabor and Wavelet Feature Extraction Techniques," *International Journal of Engineering and Innovative Technology*, vol. 2, no. 1, pp. 92-95, 2012.
- [4] Delac K. and Grgic M., *Face Recognition*, I-Tech Education and Publishing, Vienna, Austria, 2007.
- [5] Gonzalez R. and Woods R., *Digital Image Processing*, Prentice Hall, 2007.
- [6] Gonzalez R., Woods R., and Eddins S., *Digital Image Processing using MATLAB*, Gatesmark Publishing, 2009.
- [7] Gu H., Guangda Su., and Du C., "Feature Points Extraction from Faces," available at: http://sprg.massey.ac.nz/ivcnz/Proceedings/IVC NZ_28.pdf, last visited 2003.
- [8] Haykin S., *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ, 1999.
- [9] Hu X., "Affective Recognition Based on Gabor Wavelet Transformation and Fractal Dimension," *Changsha University of Science and Technology*, 2008.
- [10] Ioannou S., Raouzaiou A., Tzouvaras V., Mailis T., Karpouzis K., and Kollias S., "Emotion Recognition through Facial Expression Analysis based on a Neuro Fuzzy Network," *Journal Neural Networks*," vol. 18, no. 4, pp. 423-435, 2005.
- [11] Khandait P., Thool R., and Khandai D., "Comparative Analysis of ANFIS and NN Approach for Expression Recognition using Geometry Method," *Journal Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 3, pp. 169-174, 2012.
- [12] Lin K., Cheng W., and Li J., "Facial Expression Recognition based on Geometric Features and Geodesic Distance," *International Journal of Signal Processing*, vol. 7, no. 1, pp. 323-330, 2014.
- [13] Murtaza M., Sharif M., Raza M., and Shah J., "Analysis of Face Recognition Under Varying Facial Expression: A Survey," *the International Arab Journal of Information Technology*, vol. 10, no. 4, pp. 378-388, 2013.
- [14] Pantic M. and Patras I., "Dynamics of Facial Expression: Recognition of Facial Actions and

Their Temporal Segments from Face Profile Image Sequences," *IEEE Transactions on Systems, Man and Cybernetics-Part B: cybernetics*, vol. 36, no. 2, pp. 433-449, 2006.

- [15] Saeed K. and Charkiewicz C., *An Experimental Criterion for Face Classification*, Springer, 2003.
- [16] Sandbach G., Zafeiriou S., Pantic M., and Rueckert D., "Recognition of 3D Facial Expression Dynamics," *Journal of Image and Vision Computing*, vol. 30, no. 10, pp. 762-773, 2012.
- [17] Thai L., Nguyen N., and Hai T., "A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network," *International Journal of Machine Learning and Computing*, vol. 1, no. 4, pp. 388-393, 2011.
- [18] www.wikipedia.org, available at: https://en.wikipedia.org/wiki/Facial_expression, last visited 2014.
- [19] Zeng J. and Li D., "An Adaptive Canny Edge Detector using Histogram Concavity Analysis," *International Journal of Digital Content Technology and its Applications*, vol. 5, no. 6, pp.172-181, 2011.
- [20] Zhang S., Zhao X., and Lei B., "Facial Expression Recognition Based on Local Binary Patterns and Local Fisher Discriminant Analysis," *Wseas Transactions on Signal Processing*, vol. 8, no. 1, pp. 21-31, 2012.
- [21] Zhang Z., "Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiments with a Multi-Layer Perceptron," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 7, pp. 757-763, 1998.
- [22] Zhou J., Xu T., and Gan J., "Feature Extraction based on Local Directional Pattern with SVM Decision-level Fusion for Facial Expression Recognition," *International Journal of Bio-Science and Bio-Technology*, vol. 5, no. 2, pp. 101-110, 2013.



Ahmad Tayyar obtained Mathematics from Aleppo University, Syria in 1981; MSc and PhD in Mathematics and Informatics from FST Faculty, Hute Alsace University, France in 1987 and 1991 respectively. He is a Professor at

Computer Science Department at the Faculty of Information Technology, Jerash University, Jordan. His research interests include biometrics, computer vision, image processing and artificial intelligence.



Shadi Al Shehabi graduated from Aleppo University, faculty of science, department of Mathematics/ Computer Science in 1997. He received his MSc and PhD in Computer Science (Artificial Intelligence), University of Henry

Poincaré, France in 2002 and 2006, respectively. He is Assistant Professor in Türk Hava Kurumu Üniversitesi, Turkey. His research interests include artificial intelligence, data mining, data analysis, information retrieval and data visualization.



Majida AlBakoor obtained Mathematics from Aleppo University, Syria in 1995; MSc in Informatics Mathematics from Aleppo University, Syria in 2007. Now is PhD student in Computer Science at Mathematics Department,

Aleppo University, Syria. His research interests include biometrics, computer vision, image processing and artificial intelligence.