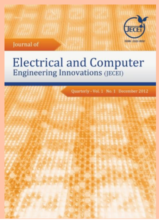
http://jecei.srttu.edu

SRTTU

Journal of Electrical and Computer Engineering Innovations

JECEI, Vol. 2, No. 2, 2014

Regular Paper



Facial Expression Recognition Based on Anatomical Structure of Human Face

Sina Mohseni[1,*], Gholamreza Ardeshir[2], and Niloofar Zarei[3]

[1,2]Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology

[3]Faculty of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran

*Corresponding Author' Information: Sina.Mohseni89@gmail.com

ARTICLE INFO

ARTICLE HISTORY:
Received 28 May 2014
Revised 7 June 2014
Accepted 30 June 2014

KEYWORDS:
Facial expression analysis
Human facial anatomy
Adaboost classifier
Support vector machine

ABSTRACT

Automatic analysis of human facial expressions is one of the challenging problems in machine vision systems. It has many applications in human-computer interactions such as, social signal processing, social robots, deceit detection, interactive video and behavior monitoring. In this paper, we develop a new method for automatic facial expression recognition based on facial muscle anatomy and human face structure. The algorithm finds approximate location of effective facial muscles and extracts features by measuring skin texture in 11 local patches. Seven facial expressions, including neutral are being classified in this study using AdaBoost classifier and other classifiers on MMI databases. Experimental results show that analyzing skin texture from selected local patches gives accurate and efficient information in order to identify different facial expressions.

1. Introduction

Emotions are often conveyed through the gestures of face and body rather than verbal communication. A pronounced feeling may be countered by a facial expression, emphasizing the feeling conveyed by the expression through sarcasm. Mehrabian [1] points out an irrefutable example; it is possible to verbally express hatred and pass on exactly the opposite feeling. In his later works, he goes one step further and claims that the feeling or attitude conveyed by a speaker is 55% facial, 38% vocal and only 7% verbal [2].

Until recently, the task of facial expression analysis has been a topic of research primarily associated with the field of psychology. However, automatic facial expression analysis has attracted much attention in the field of computer science. Automatic facial expression recognition plays a significant role in human computer interaction systems, Robotics, machine vision, virtual reality, user profiling, broadcasting, web services, border security systems, health support appliance, monitoring of stress and fatigue [3]-[5]. Recently, significant advances have been made in the area of face recognition and facial expression recognition [6]-[9]. However, there are still many challenges remaining. For example, face recognition in uncontrolled environments and conditions is still limited by lighting, head angle and the person's identity [10].

Human faces are extremely similar, thus the extraction of facial features and selection of an appropriate classifier are the two key steps to solve the facial expression recognition problems. Two main methods of feature extraction in the current research are texture-based analysis (e.g. pixel intensity) and geometry-based analysis (e.g. movement detection). In geometry-based methods, image filters, such as Gabor wavelets, are applied to either the whole face or specific face regions, called patches, to extract the appearance changes of the face. One of the most frequently used texture-based feature extraction method is Gabor filter bank [11]-[13], however, it is both time and memory intensive to convolve face images with a bank of Gabor filters to extract multi-scale and multi-orientational coefficients. Local Binary Pattern (LBP) [14]-[16] features were proposed originally for texture analysis, and recently have been introduced to represent faces in facial images analysis. The most important properties of LBP features are their tolerance against illumination changes and their computational simplicity.

On the other hand, geometry-based methods extract information using shape and location of facial components and use this information to form feature vectors [17]-[20]. The problem with implementations of geometric methods is that they usually require accurate and reliable facial point detection and tracking, which is difficult in many real world applications. Generally, we have found all geometry based methods for facial expression recognition semi-automatic and dependent on landmarks. The results are less accurate for images of subjects with blond hair or black skin. Moreover, previous methods are sensitive to head angle and pose which is inevitable in a sequence of facial expression images.

In this work, we empirically study facial representation based on local binary pattern features [21]-[22] for person-independent facial expression recognition. Our main contributions are briefly stated as follows:

1- Proposing a robust face patching method which selects effective face are as andextracts accurate information.
2- We examine three machine learning methods, including Naïve Bayes, Support Vector Machine (SVM) and AdaBoost to perform facial expression recognition using LBP features.

The rest of this paper is organized as follows. Section 2, reviews facial expression muscle anatomy. Section 3, describes proposed face batches for further texture analysis. Section 4, explains LBP feature extraction method. Section 5, represents and explains classifiers used in this research. Section 6, illustrates experimental resultsand bestrecognition rates achieved. Finally, Section 7 concludes the paper along with some suggestions for future developments.

2. A Review on Facial Expression Anatomy

The facial muscles work differently in compare is on with the other skeletal muscles because they move skin instead of joints. The muscles of facial expression are located in the subcutaneous tissue, originating from bone and inserting onto the skin. By shrinking, the muscles pull on the skin and exert their effects. They are the only group of muscles that insert into skin. All the muscles of facial expression are innervated by the facial nerve. The facial muscles can broadly be split into three groups; orbital, nasal and oral [23].

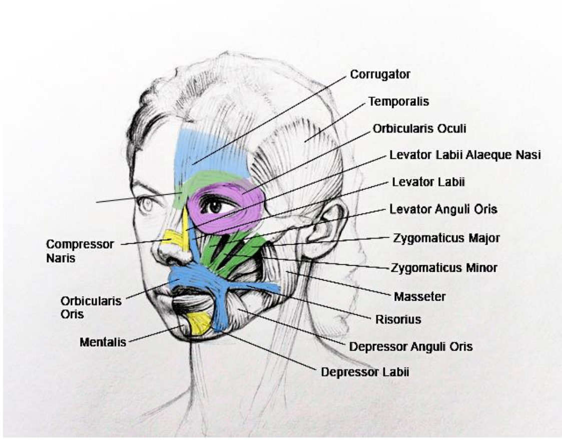


Figure 1: Facial expression muscle anatomy [24]

A. *Orbital Group*

The orbital group of facial muscles contains two muscles associated with theeyesocket. These muscles control the movements of theeyelids, and eyebrows. Orbicular is Oculisur rounds the eye socket and extends into the eyelid, performing closure of the eyelid. The Corrugator Supercilii is a much smaller muscle and inserts into the skin of the eyebrow.It acts to draw the eyebrows together, creating vertical wrinkles on the bridge of the nose [24].

B. *Nasal Group*

The nasal group of facial muscles are associated with movements of the nose, and the skin around it. The nasalis is the largest of the nasal muscles. It is split into two parts which have opposing functions. The upper part compresses the nares, and the lower part opens the nares. The procerus is the most superior of the nasal muscles. Contraction of this muscle pulls the eyebrows downward to produce transverse wrinkles over the nose [24].

C. *Oral Group*

These are the most important group of the facial expressors. They are responsible for movements of the mouth and lips. Such movements are required in singing and whistling, and add emphasis to vocal communication. The Orbicularis Ores enclose the opening to the oral cavity and purses the lips. The Buccinators pulls the cheek inwards against the teeth, preventing accumulation of food in that area. There are other muscles that act of the lips and mouth, anatomically, they can be divided into upper and lower groups. The lower group contains the depressor anguli oris, depressor labii inferioris and the mentalis. The upper group contains the risorius, zygomaticus major, zygomaticus minor, levatorlabii superioris, levatorlabii superior is a laequenasi and levatoran gulioris

[24].

3. Affective Muscles and Skin Regions

Facial muscles do not only regulate the position and width of facial openings but also make them more expressive. By those means, the face is able to convey emotions and the present psychological state of a person, which plays an extraordinary role in the nonverbal communication between people. Studies in facial expression anatomy show that any movement in facial skin is because of facial muscle shrinkage.

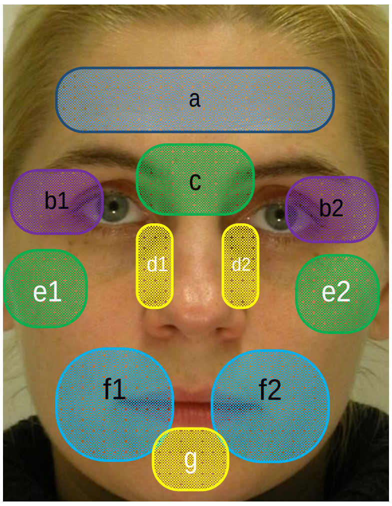


Figure 2: Selected facial regions for feature extraction based on facial muscle anatomy

Facial muscles are links between bones and skin tissues. Thus, skin deformation or shrinkage happen somewhere in the middle of muscle length.

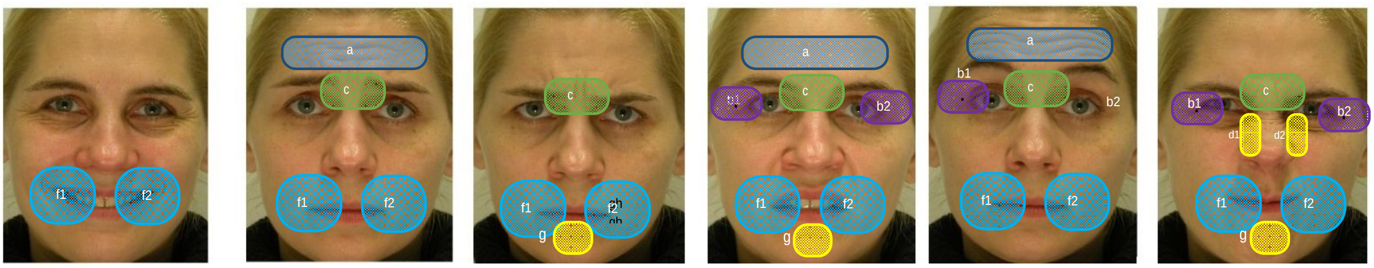


Figure 3: Active action units (AUs) for each of seven facial expressions are recognized in local patches areas

Visual changes and deformations on skin can be captured and analyzed in robotics and other machine vision systems. Therefore, 11 local patches are located around corresponding effective muscles in order to extract maximum texture features from the skin surface, as shown in figure 2. Below is a list of selected patches and their covered facial muscles:

a) Inner/outer corner of eyebrow raised.
b) Eyes narrowed/widened.
c) Eyebrows drawn medially and down.
d) Upper lip raise and inverted.
e) Cheeks raised; eyes narrowed.
f) Lip corners pulled up; Angle of the mouth elevated; Lip corners tightened; Lips tightened; Lips pressed together; Jaw dropped.
g) Skin of chin elevated; Lower lip pulled down and laterally.

Groups of facial muscles movements can lead to various facial action coding system (FACS) action units (AUs). Although Ekman and Friesen proposed that specific combinations of FACS action units represent prototypic expressions of EMFACS emotions [25]. FACS itself is purely descriptive and includes no inferential labels. By converting FACS codes to EMFACS or similar systems, face images may be coded for 7 main emotion-specified expressions.

Figure 3 illustrates six main facial emotions which are analyzed for active AUs. Table 1 gives relative information about various AUs and their related muscle movements. It can be seen that there are 8 main AUs forming emotional expressions. Column three shows selected affective local patches for action units tracking.

4. Local Binary Pattern (LBP)

The most important property of LBP features is their tolerance against illumination changes and their computational simplicity.

TABLE 1
ACTION UNITS FORMING EMOTIONAL EXPRESSIONS AND AFFECTIVE LOCAL PATCHES FOR EACH ACTION UNIT SET

Emotional Expression	Action Units	Local Patches
Happiness	6 + 12	e + f
Sadness	1 + 4 + 11 +15	a + c + f
Anger	4 + 5 + 7 + 10 + 22 + 23 +26	c + f + g
Fear	1 + 2 + 4 + 5 + 20 + 25	a + b + c + f + g
Surprise	1 + 2 + 5 + 26 + 27	a + b + c + f
Disgust	9 + 15 + 16 + 17 + 26	b + c + d + f + g

In primary LBP operator, it labels the pixelsimage by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number, see Fig. 4 for an illustration. Then, the histogram of the labels can be used as a texture descriptor. The derived binary numbers, called Local Binary Patterns codes, codify local primitives including different types of curved edges, spots and flat areas over the whole image, so can be used to statistically describe image characteristics, Fig. 4.b.

The limitation of the basic LBP operator is its small 3×3 neighborhood which cannot capture dominant features with large scale structures. Later, the LBP operator is extended to use neighborhood of different sizes. Using circular neighborhoods and bilinear interpolating the pixel values allow any radius and number of pixels in the neighborhood.

The LBP operator with P sampling points on a circular neighborhood of radius R is given by:

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(g_p - g_c)2^p \quad (1)$$

where, g_c is the gray value of the central pixel, g_p is the value of its neighbors, P is the total number of involved neighbors and R is the radius of the neighborhood.

5. PROPOSED RECOGNITION TECHNIQUE

Feature selection along with the regions from where these features are going to be extracted is one of the most important steps to recognize expressions. As the proposed framework draws its inspiration from the human muscle anatomy, it extracts LBP features, only from the perceptual salient facial regions which were determined through anatomy-visual experiment.

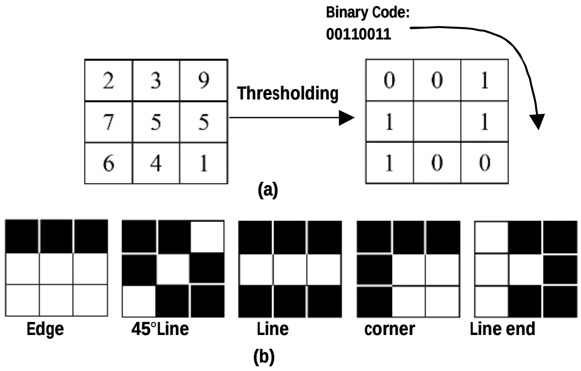


Figure 4: (a) The basic LBP operator in a sample 3×3 block. (b) Examples of texture primatives which can be detected by LBP. White and black pixles represents ones and zeros.

Steps of the proposed framework are as follows, the framework first localizes face region using Viola–Jones face and facial object detection algorithm [26]. We selected this algorithm as it is the most cited and is considered the most accurate pattern recognition method for face and facial region detection. Next, the face image is sub-divided into 11 selected regions. We localize selected patches in the area of detected face using viola & jones object detection tool, Fig. 5 illustrates patches with yellow boxes. Then, the framework extracts LBP features from selected patches. At last, the classification is carried out on the basis of extracted features in order to recognize action units in each selected patch. Occurrence of AUs together, according to table 1, results in various facial expression recognition.

TABLE 2
CLASSIFICATION ACCURACY (%) IN DIFFERENT CLASSIFIERS WITH CROSS-VALIDATION

Expression	Surprise	Happiness	Disgust	Fear	Anger	Sadness	Neutral
SVM	74.2	75.3	71.6	78.4	71	69.5	100
Naive Bayes	67	71.3	59.5	54.5	62.8	62.6	98.5
Adaboost	80.9	81.7	75.6	72.8	74.2	69.4	99.5

TABLE 3
CONFUSION MATRIX (%) FOR THE BEST RESULT OF ADABOOST CLASSIFIER

	Surprise	Happiness	Disgust	Fear	Anger	Sadness
Surprise	80.9	6.3	12.7	0	0	0
Happy	5.3	81.7	8.2	4.8	0	0
Disgust	7.6	3.3	75.6	7.3	6.2	0
Fear	4.2	0	4.1	72.8	6.1	12.8
Angry	0	7	3.4	6.7	74.2	8.7
Sad	5.4	0	2.9	9.5	12.8	69.4

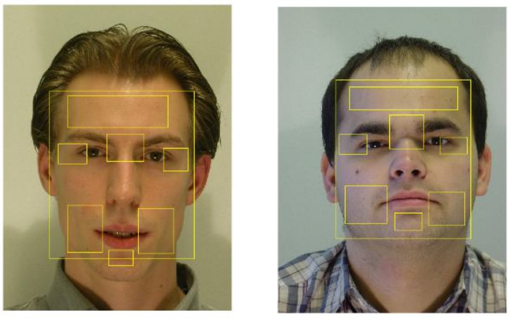


Figure5: Example of salient patches detected in the face image

6. EXPERIMENTAL RESULTS

All the results are obtained using MATLAB software. We performed subject-independent facial expression recognition using proposed anatomical facial expression analysis. In order to verify the proposed method, we used MMI database [27]. This data set is publicly available and has been used for several other publications. This facial expression database is recorded in true color with a frame rate of 24 fps. The advantage of using this database is that it contains a large number of videos that display facial expressions with a neutral-apex-neutral evolution. The video sequences chosen for inclusion in the data were based on the attached label for the six basic expressions.

In this step of experiment, we only select natural and apex frames. 19 frequent AUs were selected and 256 dimension feature vectors were extracted from each facial patch. We used classifiers for each patch and 6 main expressions (happiness, surprise, disgust, sadness, fear, anger) and neutral are classified from action unit results. Since we used 10 fold cross validation in both classifiers, we split each class of expression by a ratio of 10/1 for training/testing. To eliminate variations among results, 20 rounds of validation are performed using generated splits, and the final result is calculated by averaging over all rounds.

Happiness and surprise with salient features such as open mouth can be recognized effortlessly. On the contrary, sadness and anger are much less distinguishable. Table 2 shows the per-class recognitions accuracy of different methods. As can be seen in Table 2, tuned SVM and AdaBoost give about the same results for recognizing posed expressions. While Naïve Bayes shows relatively lower accuracy. What is noteworthy is that ensemble classification methods achieve comparable performance to other individual classifiers, because of their weighted voting nature. As an example, here we only give the confusion matrix obtained from AdaBoost classifier in Table 3.

7. CONCLUSION

We presented a novel descriptor for reliable facial action units and facial expression recognition. Framework is based on facial muscles anatomy and works adequately on posed and spontaneous expressions. Anatomy based selected facial regions used in this research consists of 11 patches which extract LBP information from skin region. Experiments were done on three different classifiers to demonstrate the efficiency of proposed method; NB, SVM and AdaBoost. The best performance on the classification of six basic expressions and neutral was 79.1%, obtained on MMI database through AdaBoost. It should be noted that, the results are not directly comparable due to different databases and experimental setups.

The key conclusions drawn from this study are:

1- Facial expressions can be analyzed automatically by extracting features from salient facial regions.

2- In proposed face image analysis, we extract features from 11 specified regions to detect occurrence of certain action units and then system finds nearest emotion category.

3- Features extracted from facial skin patches have strong discriminative ability as the recognition results for seven universal expressions is not affected so much by the choice of classifier.

In the future, we plan to test the proposed framework on other famous databases, also working with video data gives more realistic results. Furthermore, summation of audio and visual data to understand human emotional situation is an interesting topic for future investigations.

REFERENCES

[1] A. Mehrabian, Silent Messages: Implicit Communication of Emotions and Attitudes, Wadsworth Pub. Co., 1981.

[2] A. Mehrabian, Communication without words, Psychol. Today 2 (9) ,1968, 52–55.

[3] B. Golomb, T. Sejnowski, "Benefits of machine understanding of facial expression" in NSF Report- Facial expression Understanding, Salt Lack City, UT, 1997, pp 55-71.

[4] M. Pantic, "Face of ambient interface", in Ambient Intelligence in Everyday Life, vol. 3864, lecture notes on Artificial Intelligence. Berlin, Germany: Springer-Verlag, 2006, pp. 32-66.

[5] A. Young, Face and Mind, Oxford, U.K., Oxford Univ. Press, 1998.

[6] Shaohua Wan n, J.K.Aggarwal: Spontaneous facial expression recognition: A robust metric learning approach. J. Pattern Recognition, 47, 1859–1868, 2014.

[7] M. Taner Eskil, Kristin S. Benli: Facial expression recognition based on anatomy. J. Computer Vision and Image Understanding, 119, 1–14, 2014.

[8] Yongqiang Li, Shangfei Wang, Yongping Zhao, and Qiang Ji: Simultaneous Facial Feature Tracking and Facial Expression Recognition, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 7, 2013.

[9] Michel F. Valstar, Maja Pantic,: Fully Automatic Recognition of the Temporal Phases of Facial Actions, IEEE Transactions on systems, man, and cybernetics, vol. 42, no. 1, 2012.

[10] B. Fasel, J. Luettin, Automatic facial expression analyses: a survey, Pattern Recognition, 36, 259-75, 2003.

[11] M. S. Bartlett, G. Littlewort, M. G. Frank, C. Lainscsek, I. Fasel, and J. R. Movellan, "Recognizing facial expression: Machine learning and application to spontaneous behavior," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 2, 568–573, 2005.

[12] M.S. Bartlett, G.C. Littlewort, M.G. Frank, C. Lainscsek, I.R. Fasel, and J.R. Movellan, "Automatic Recognition of Facial Actions in Spontaneous Expressions," J. Multimedia, vol. 1, no. 6, pp. 22-35, 2006.

[13] W.Gu, C.Xiang, Y.V.Venkatesh, D.Huang, H.Lin, Facial expression recognition using radial encoding of local Gabor features andc lassifier synthesis,Pattern Recognition, 45, 80–91, 2012.

[14] G. Zhao, M. Pietikäinen, Boosted multi-resolutions patiotemporal descriptors for facial expression recognition, Pattern Recognition Letters, 30 (no.12), 1117–1127, 2009.

[15] C. Shan, S. Gong, P.W. McOwan, Robust facial expression recognition using local binary patterns, in: Proceedings of the IEEE International Conference on Image Procession, 370-373, 2005.

[16] Yang, P., Liu, Q., Metaxas, D.N., Exploring facial expressions with compositional features. In: IEEE Conference on Computer Vision and Pattern Recognition, 2010.

[17] M. Pantic, L.J. Rothkrantz, Facial action recognition for facial expression analysis from static face images, Trans. Syst. Man Cyber. Part B 34 (3), 1449–1461, 2004.

[18] S. Lucey, I. Matthews, Ch. Hu, Z. Ambadar,: AAM Derived Face Representations for Robust Facial Action Recognition, In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, 2006.

[19] Y. Zhang, Q. Ji, Active and dynamic information fusion for facial expression understanding from image sequences, IEEE Trans. Pattern Anal. Mach. Intell. 27 (5), 699–714, 2005.

[20] Hui Fang, N. M. Parthaláin, A. J. Aubrey, G. K.L. Tama, R. Borgo, P. L.Rosin, P. W. Grant, D. Marshall, M. Chen,: Facial expression recognition in dynamic sequences: An integrated approach, Pattern Recognition, 47, 1271–1281, 2014.

[21] T. Ojala, M. Pietikäinen, T. Mäenpää: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7), 971–987, 2002.

[22] C. Shan, Sh. Gong, P. W. McOwan: Facial expression recognition based on local binary patterns: a comprehensive study, Image and video computing 27, 803-816, 2009.

[23] Gassner HG, Rafi i A, Young A, Murakami C, Moe K, Larrabee WF (2008) Surgical anatomy of the face. Implications for modern face-lift techniques. Arch Facial Plast Surg 10(1):9–19.

[24] M. J. Fehrenbach, S. W. Herring, Illustrated Anatomy of the Head and Neck, 4th Edition.

[25] W.V. Friesen, P. Ekman, EMFACS-7: Emotional Facial Action Coding System, Unpublished manuscript, University of California at San Francisco, 1983.

[26] P. Viola and M. Jones, "Robust real-time object detection," Int. J. Comput.Vis.., vol. 57, no. 2, pp. 137–154, 2004.

[27] MMI Face Database. Availible online: Http://www.mmifacedb.eu/

[28] M.F. Valstar, M. Pantic, "Induced Disgust, Happiness and Surprise: an Addition to the MMI Facial Expression Database", Proceedings of the International Language Resources and Evaluation Conference, Malta, May 2010

[29] M. Pantic, M.F. Valstar, R. Rademaker and L. Maat, "Webbased database for facial expression analysis", Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME'05), Amsterdam, The Netherlands, July 2005

BIOGRAPHIES

Gholamreza Ardeshir received his B.Sc. from Ferdosi University (Mashhad-Iran) in 1989, M.Sc. from Tarbiat Modares University (Tehran-Iran) in 1993 and Ph.D. from University of Surrey (Guildford-UK) in 2003, respectively. Since 1994, He has been a member of Electrical Engineering Faculty at Babol University of Technology. His research interests are VLSI circuit design and signal processing.

Sina Mohseni was born in Babol, Iran, in 1989. He received the B.Sc. degree in electronic engineering from Isfahan University, Isfahan, Iran, in 2011. He finished his M.Sc. degree in electronic engineering with focus on computer vision systems in 2013. Since 2013, he has been an assistant and researcher with the Signal Processing Laboratory, Babol Noshirvani University of Technology. His scientific fields of interest include face recognition, facial expression recognition and human pose estimation.

Niloofar Zarei was born in Tehran, Iran, in 1991. She received her B.Sc. degree in electronic engineering from Amirkabir University of Technology, Tehran, Iran, in 2014. Her scientific fields of interest include face recognition, facial expression recognition and human pose estimation.

J. Elec. Comput. Eng. Innov. 2014, Vol. 2, No. 2, pp. 77-83