6th International Conference
ðComputational Mechanics and Virtual Engineering ð
COMEC 2015
15-16 October 2015, Braşov, Romania

DETECTING HUMAN EMOTIONS WITH AN ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

Phd. Gheorghe Gîlcă[1], Proffesor Nicu-George Bîzdoacă[1]
[1] University of Craiova, Craiova, ROMANIA, gigi@robotics.ucv.ro, nicu@robotics.ucv.ro

Abstract: *This article deals with an emotion recognition system based on fuzzy sets. Human faces are detected in images with the Viola - Jones algorithm and for their tracking in video sequences we used the Camshift algorithm. The detected human faces are transferred to the sugeno type decisional fuzzy system, which is based on the variable fuzzyfication measurements of the face: eyebrow, eyelid and mouth. The system can easily determine the emotional state of a person.*
Keywords: *emotion, face detection, face tracking, fuzzy rules, training*

1. INTRODUCTION

The verbal and nonverbal communication is the most common in relationships between people, respectively between machines and people. Mehrabian [1] explained in his work in what percentage each form of communication contributes to the overall message delivered between two individuals: the linguistic message is only 7%, the vocal intonation is 38% while the facial expression is 55% of the sent message .
The recognition and classification of the human facial expressions is an important aspect to develop in an automatic recognition system in order to facilitate the communication between machines and people.
Facial emotions are used in areas such as telemedicine, distance learning, robotics and automotive. There are six main types of basic emotions: happiness, anger, fear, sadness, disgust and surprised. All other emotions are variations of this basic emotions. Each emotion is characterized by psychological and behavioral qualities, including movement, posture, voice, facial expression and heart rate fluctuation. Although there are many different types of emotions, they all have some common characteristics.
Charles Darwin was the first to bring into discussion the universality of emotions, using this idea in his book [2], to support the theory of evolution, arguing that emotions are mental reaction patterns imprinted in the nervous system.
The American psychologist Paul Ekman from the University of California (expert in the study of emotions, their manifestations and the study of the mechanism of lies), based on the results of a long research on all the continents, finds that there are six facial expressions which can be recognized by any person belonging to any culture on the planet: fear, anger, sadness, surprise, disgust and joy [3]. The universality of these events may be considered to be a strong indication that these emotions are six basic emotions related to human nature.
Adaptive Neuro-Fuzzy Inference System (ANFIS) classifies the expressions of a supplied face into seven basic categories like: surprised, neutral, sad, disgusted, fear, happy and angry.

2. RELATED WORKS

Fatemeh Shahrabi Farahani et al., present a new method based on fuzzy logic for recognizing emotion based on the eyes and mouth features. Their approach shows an emotion recognition system which has 3 levels: face detection, feature extraction and classification. The fuzzy logic used by the authors in the third stage is performed using the Mamdani type inference relations with 94 rules of the trapezoidal membership functions to encode facial attributes and their mapping of the emotion space [4].
Joseph W. Matiko et al. present in their paper, an algorithm designed to classify emotions into two categories: positive and negative. Their proposed algorithm is based on fuzzy sets making the fuzzyfication of signal from

the EEG. Their method is advantageous compared to the literature, because the classification includes the type of emotions, but also its power [5].
In their paper, Yong-Hwan Lee et al. proposed a method of extraction and recognition of facial expression and emotions on mobile cameras. They formulated a classification model using 65 landmarks in order to estimate facial expressions [6]. With their method we can recognize three types of emotions: neutral, happy and angry. Khandait S.P. et al. perform in their work a comparative study of two methods: The Fuzzy Inference Adaptive Neuro-System (ANFIS) and The Back Propagation Neural Network (BPNN) to classify facial expressions [7].

3. METODOLOGY

We have proposed an emotions recognition system with a simple structure as in Figure 1.
The system consists of three subsystems: the first is to detect the human face in images, the second is to track the human face, and the third is to decide what emotion is shown on the human face.

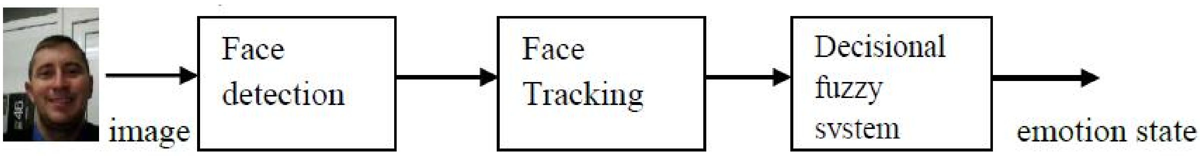


Figure 1: Structure of proposed system

3.1. Face detection using the Viola-Jones algorithm

Figure 2 [8] illustrates the Viola and Jones object detection algorithm in two steps: the detecting of the face in an image region and then applying on it a cascade of boosted classifiers. The first step detection window is realized by scanning the same image many times, each time with a new size. During the second stage each window is passed through a cascade of classifiers which in its turn is divided in various steps, each step uses a set of weak learners.
Each stage is trained to select only the wanted images using a technique called boosting. Boosting has the advantage of training a very accurate classifier to choose a weighted average of the decisions taken by the weak learners. Each level of the labeled region belonging to the classifier is defined as being the current location of the sliding window, being positive or negative.
The detector reports a found face in the current location when the final level classifies the region as being positive. If we have the k classifiers in a cascade, the result of detection rate, D, and false positive rate, F, is given by the product rates on each stage classifier [9]:

$$D = \prod_{i=1}^{k} d_i \,, \quad F = \prod_{i=1}^{k} f_i \qquad (1)$$

,where d_i is the detection rate of the *i* the classifier in the examples that get through to it and f_i is the false detection rate of the *i* the classifier in the examples that get through to it.

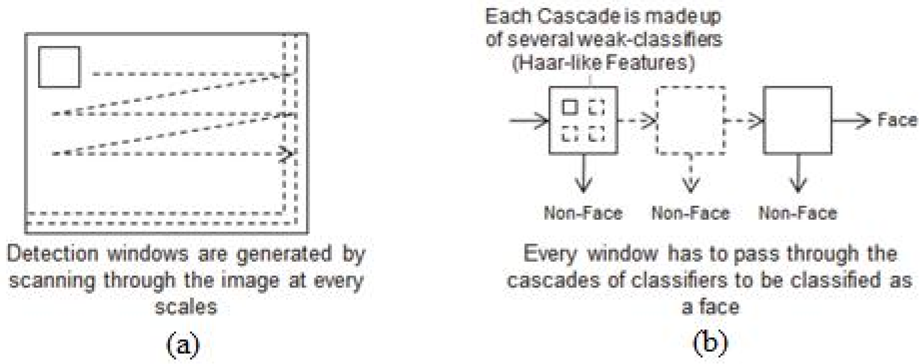


Figure 2: The Viola-Jones Object Detection Algorithm: a) Detection windows ; b) Classified windows

3.2. Face tracking using the Camshift algorithm

For face tracking in a video sequence, we used the camshift algorithm proposed by the authors in their paper [10]:

a) Set the calculation region of the probability distribution for images.
b) Choose an initial location of the Mean Shift search window; it will be the tracked target.
c) Calculate a color probability distribution of the region centered on the Mean Shift search window, where the target is slightly larger than Mean Shift window size.
d) Run Mean Shift algorithm to find the centroid of the probability image. Store the zero moment (the area) and the centroid location.
e) For the next frame, put the search window at the mean location found in Step d) and set the window size to a function of the zeroth moment. Go to Step c).

3.3. The decisional fuzzy system

After the operations of detection and tracking of the human face, it goes through a fuzzy system modeled using the Fuzzy Logic Toolbox of Matlab. The system takes into account in making the decision three face variables that change for each emotional state. Our system is modeled to recognize seven states: happiness, anger, fear, sadness, disgust, surprise and neutral. More details of this system are given in the next section.

4. THE DESIGN OF THE ADAPTIVE NEURO-FUZZY LOGIC SYSTEM IN DETERMINING EMOTION RECOGNITION

The fuzzy model projected by us looks like Figure 3:

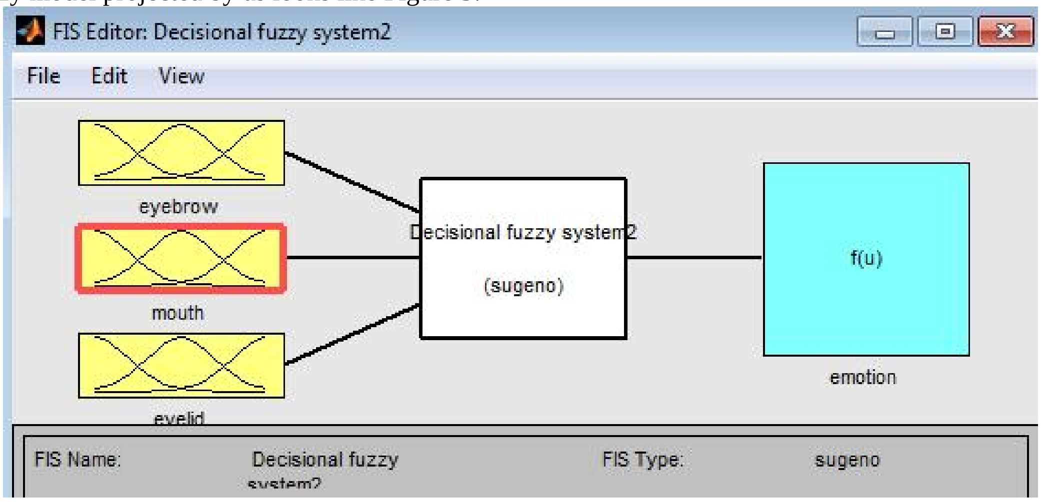


Figure 3: The fuzzy system model

The three yellow rectangles represent the three inputs of the fuzzy logic system: eyelid, eyebrow and mouth. In these three blocks the fuzzyfication takes place. The light gray rectangle represents the process of inference and the turquoise rectangle represents the output fuzzyfication block (in these rectangle the defuzzyfication takes place).

4.1. Inputs

In this subsection we define universes of discussion for the 3 inputs variables thus:

- Eyelid: [40; -100];
- Eyebrow: [100;-100]:
- Mouth: [10;-100].

The associated fuzzy set of input variables are: small, moderate and large.

4.2. Outputs

For the output variable emotion, the universe of discussion is [0;1.4] and the associated fuzzy sets are: happy, sad, surprised, fear, anger, disgusted and neutral, each with the following universes of discussion, in order: [0;0.2], [0.2;0.4], [0.4;0.6], [0.6;0.8], [0.8;1], [1;1.2], [1.2;1.4].
The Sugeno linear output model has the general form:

$$a_0 X_0 + a_1 X_1 + \ldots + a_i X_i + z \quad (2)$$

Where a_i is a constant parameter and X_i is input variables. For constant functions, the values of a_i is equal to zero and, hence Y = z (constant value). For the linear function model, the a_i values are entered through the ANFIS editor, while X_i are the values of the input factors. The correlation between the input and the output variables is done through a set of fuzzy rules. Each rule uses AND/OR connectors to connect various input factors with a particular output emotion.

4.3. The defining of a set of rules for the decisional fuzzy system

The system works with fuzzy rules type if ..., then ..., that condition of if must be complied with in order to deduce the desired output, which makes the system to be decisional. We have created a knowledge base for the system with the seven following rules:

- If (the mouth is moderate) and (Eyelid is small) and (Eyebrow is moderate) then (Emotion is happy);
- If (mouth is small) and (Eyelid is large) and (Eyebrow is moderate) then (Emotion is sad);
- If (Mouth is moderate) and (Eyelid is large) and (Eyebrow is small) then (Emotion is anger);
- If (Mouth is large) and (Eyelid is large) and (Eyebrow is large) then (Emotion is surprised);
- If (Mouth is moderate) and (Eyelid is large) and (Eyebrow is large) then (Emotion is fear);
- If (Mouth is small) and (Eyelid is moderate) and (Eyebrow is small) then (Emotion is disgust);
- If (Mouth is small) and (Eyelid is moderate) and (Eyebrow is moderate) then (Emotion is neutral);

E.g. if we have the following values of the input variables: Mouth is -45, Eyelid is -100, and Eyebrow is 0 then the output will be happy.

4.4. The results design of the fuzzy system decisional system

Figure 4 shows the operation of the fuzzy system designed by us for the example above. The results are found by the system as being valid only for the first rule, where the values of the three variables belong to the fuzzy sets in the fuzzification process. The system output is that the variable has the value equal with 0.2, so it was classified as happy.

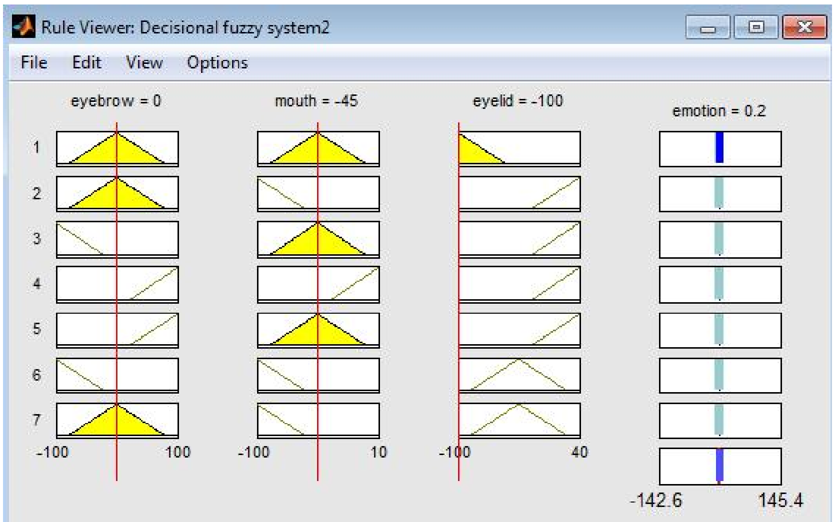


Figure 4: Happy emotional state classification by the decisional fuzzy system

4.5. The ANFIS Model Structure

The neuro-fuzzy model structure consists of five layers: the first layer is the input variables (mouth, eyebrow, eyelid), the second layer is the membership functions of inputs, which are the type gaussfm, gauss2mf and psigfm; the third layer is the rules system (seven in total); the fourth layer is the emotional state (also seven), and the fifth layer is the output system.

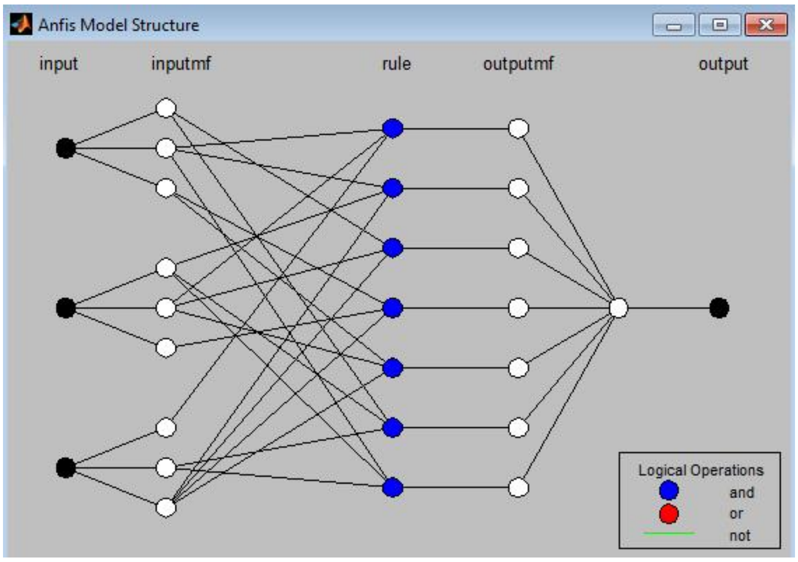
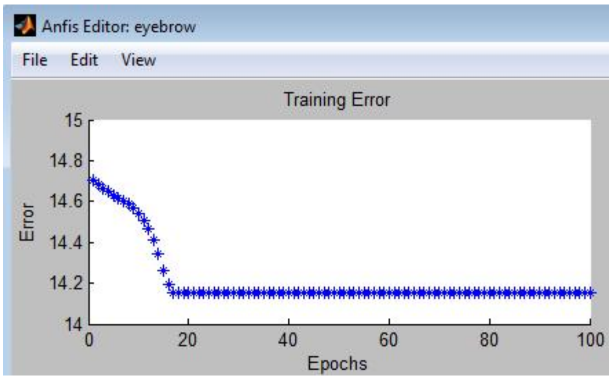
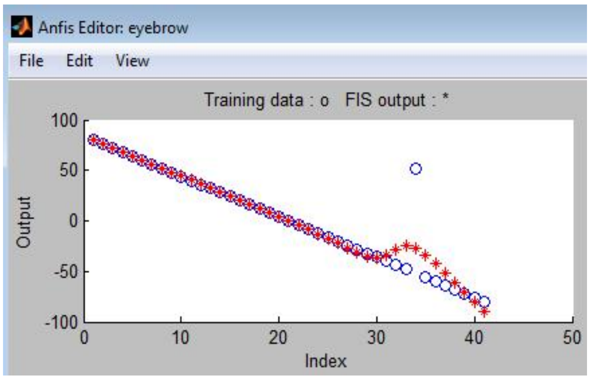


Figure 6: The Neuro-Fuzzy Structure for the proposed system

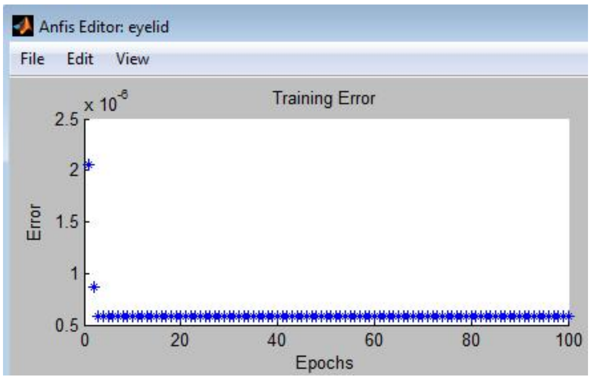
We have achieved the best results for the training of neuro-fuzzy inference system with the membership function psigmf. The results are displayed for each input in Figure 7.



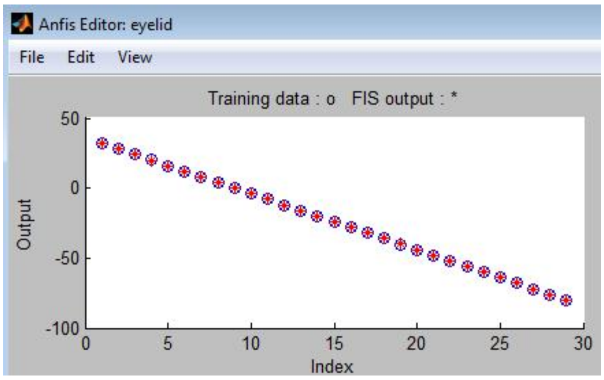
7 a)



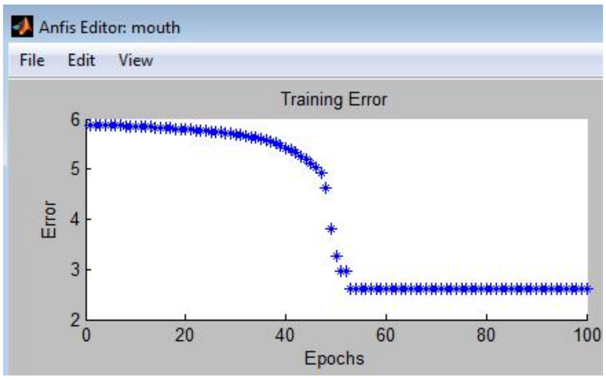
7 b)

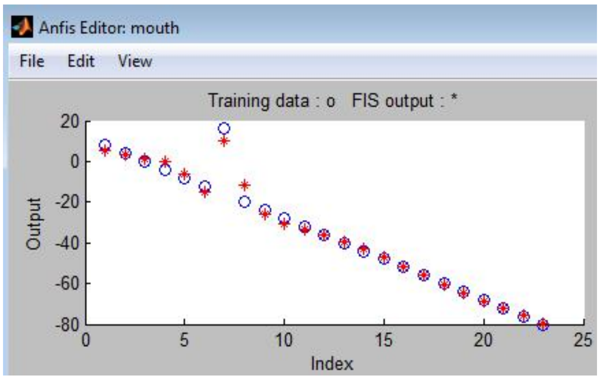


7 c)



7 d)





7 e) 7 f)

Figure 7: The experimental results for the training of neuro-fuzzy inference system

In Figure 7 a) the graph of training error of the eyebrow input variable is presented, and 7 b) shows the graph of points of its (blue circles) and the fuzzy system output (red asterisk). In Figure 7 c) the graph of training error of the eyelid input variable is presented, and 7 b) shows the graph of points of its (blue circles) and the fuzzy system output (red asterisk). In Figure 7 e) the graph of training error of the mouth input variable is presented, and 7 f) shows the graph of points of its (blue circles) and the fuzzy system output (red asterisk).

From Figure 7 we can deduce that the system has the smallest error at the eyelid input variable, which is equal with $\varepsilon=0.5x10^{-6}$ and the output data system perfectly follows the input data. For the other two variables the output system is acceptable and the training error is a little higher than the the eyebrow variable.

6. CONCLUSION

The proposed system can recognize emotions quickly and easily, because it has strict rules in its knowledge base and membership functions of the output variable are not overlapping. We presented a neuro fuzzy model of human emotion detection using three factors of the human face that can make a big emphasis in the classification of the emotional states. As future work proposed system will be expanded to recognize human emotion in a real time, which will be implemented SociBot robot.

REFERENCES

[1] Mehrabian. A, Communication without Words , Psychology Today, Vol. 2, No.4, pp 53-56, 1968.
[2] Darwin, C., The expression of the emotions in man and animals, London, John Murray, 1872.
[3] Ekman, P., Friesen, W., Facial Action Coding System: Investigator's Guide, Palo Alto, CA: Consulting Psychologists Press, 1978.
[4] Farahani, F.S., Sheikhan, M., Farrokhi A., A fuzzy approach for facial emotion recognition, Fuzzy Systems (IFSC), 2013 13th Iranian Conference on, vol., no., pp.1,4, 27-29 Aug. 2013.
[5] Matiko, J.W., Beeby, S.P., Tudor, J., Fuzzy logic based emotion classification, Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on , vol., no., pp.4389,4393, 4-9 May 2014.
[6] Lee, Y-H., Han, W., Kim, Y., Kim, B, Facial Feature Extraction Using an Active Appearance Model on the iPhone, Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2014 Eighth International Conference on , vol., no., pp.196,201, 2-4 July 2014.
[7] S.P.Khandait, R.C.Thool, P.D.Khandait, Comparative Analysis of ANFIS and NN Approach for Expression Recognition using Geometry Method, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 3, ISSN: 2277 128X, March 2012.
[8] Mustafah, Y.M., Azman, A.W., Out-of-Plane Rotated Object Detection using Patch Feature based Classifier, Procedia Engineering, Volume 41, 2012, Pages 170-174, ISSN 1877-7058.
[9] Castrillón, M., Déniz, O., Hernández, D., Lorenzo, J., A comparison of face and facial feature detectors based on the Viola–Jones general object detection framework, Volume 22, Issue 3, Pages 481-494, ISSN 0932-8092, May 2011.
[10] Allen, G.J., Xu, R.Y.D., Jesse, S.J., Object tracking using camshift algorithm and multiple quantized feature spaces, Proceedings of the Pan-Sydney area workshop on Visual information processing, Australian Computer Society, Inc., 2004.