Facial Expression Recognition based on Facial Feature and Multi Library Wavelet Neural Network

Nawel Oussaifi, Wajdi Bellil and Chokri Ben Amar
Research Group on Intelligent Machines (REGIM), National Engineering School in Sfax, Sfax University, Sfax, Tunisia

Keywords: Facial Expressions Classification, Wavelet Network, Facial Landmarking.

Abstract: In this paper, we propose a wavelet neural network-based system for automatically classifying facial expressions. This system is based on Multi Library Wavelet Neural Network (MLWNN) for emotions classification. Like other methods, our approach relies on facial deformation features. Eyes, mouth and eyebrows are identified as the critical features and their feature points are extracted to recognize the emotion. After feature extraction is performed a Multi Library Wavelet Neural Network approach is used to recognize the emotions contained within the face. This approach differs from existing work in that we define two classes of expressions: active emotions (smile, surprise and fear) and passive emotions (anger, disgust and sadness). In order to demonstrate the efficiency of the proposed system for the facial expression recognition, its performances are compared with other systems.

1 INTRODUCTION

Identifying human facial expressions has become an important study field because of its inherent intuitive appeal as well as its possible applications such as human computer interaction, face image compression, synthetic face animation and video facial image queries .This paper describes a method for facial expression recognition from image sequences using 2D appearance-based local approach for the extraction of intransient facial features, i.e. features such as eyebrows, lips, or mouth, which are always present in the image, but may be deformed .

Facial expressions are of great importance, since they usually provide a comprehensible view of users' reactions. Actually, Cohen commented on the emergence and significance of multimodality, Albeit in a slightly different human-computer interaction (HCI) domain, in (Cohen, 1998), while Oviatt (Oviatt, 1999) proposes that an interaction pattern constrained to mere "speak-and-point" only makes up for a very small fraction of all spontaneous multimodal utterances in everyday HCI (Oviatt, 1997). In the context of HCI, Jaimes in (Jaimes, 2005) defines a multimodal system as one that "responds to inputs in more than one modality or communication channel abundance". Mehrabian (1968) suggests that facial expressions and vocal intonations are the main means for someone to estimate a person's affective state (Zeng, 2007), with the face being more accurately judged, or correlating better with judgments based on full audiovisual input than on voice input (Pantic, 2003). This fact led to a number of approaches using video and audio to tackle emotion recognition in a multimodal manner (Ioannou, 2005), (De Silva, 2000), while recently the visual modality has been extended to include facial, head or body gesturing (Gunes, 2005), (Karpouzis, 2007).

2 WAVELET NEURAL NETWORK FOR EMOTION CLASSIFICATION

As it is known, the classic wavelet network CWNN has a drawback which is the low speed of convergence. For this reason new network architecture, called multi-Library wavelet neural network (MLWNN) was proposed by W. BELLIL et.al in (Chihaoui, 2010), (Bellil, 2008), (Bellil, 2007). It is similar to the classic network, but with some slight differences; the classic network uses dilation and translation versions of only one mother wavelet, but the new version constructs the network by the implementation of several mother wavelets in

Oussaifi N., Bellil W. and Ben Amar C..
Facial Expression Recognition based on Facial Feature and Multi Library Wavelet Neural Network.
DOI: 10.5220/0004034004470450
In *Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2012), pages 447-450
ISBN: 978-989-8565-22-8

the hidden layer. The architecture of MLWNN is presented in figure 1. The output of MLWNN is given in equation 1 (Othmani, 2011).

$$f(x)=\sum_{p(i,j)=1}^{N_w} w_p \Psi_p(x)+\sum_{k=0}^{N_i} a_k x_k + b \qquad (1)$$

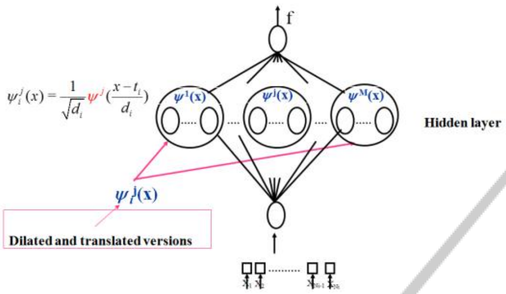

Figure 1: MLWNN architecture.

3 PRESENTATION OF THE PROPOSED APPROACH

From a physiological perspective, a facial expression results from the deformations of some facial features caused by an emotion (Zeng, 2007). Each emotion corresponds to a typical stimulation of the face muscles. The aim of our work is to evaluate the possibility of recognizing the six universal emotions by only considering the deformations of permanent facial features such as eyes, eyebrows and mouth.

The proposed algorithm is based on the analysis of permanent features calculated on the face when the emotion is product. We defined a model of each emotion that is based on the facial deformation characteristics (distance features and feature points). We propose to define two expression categories: active emotions (smile, surprise and fear) and passive emotions (anger, disgust and sadness). 2D facial emotion recognition occurs in two distinct stages: the learning phase corresponds to a representative vector of recording for each emotion of persons in a data base. The model is built from 97 faces that are essentially extracted from the base Cohn-Kanade Trade and represent different basic expressions. 21 landmarks are manually positioned on these images to generate characteristic vectors. This phase determin 97 vectors for the six emotions, then determining an average vector for each emotion. A Multi Library Wavelet Neural Network is used to approximate each vector by adjusting wavelet structural parameters (weights, dilations and translations). Six classes are generated for each facial expression. For classification we used minimum distance as defined in equation (2). Figure 2 shows facial expressions classification process.

$$\min(d_i)=\sqrt{\left(w_i^j-w_t\right)^2}\quad i=1,\dots,N_w;\ j=1,\dots,6 \qquad (2)$$

Where: N_w is the wavelet number in hidden layer

w_i^j are the weights of the class j

w_t are the weights of the 2D test face.

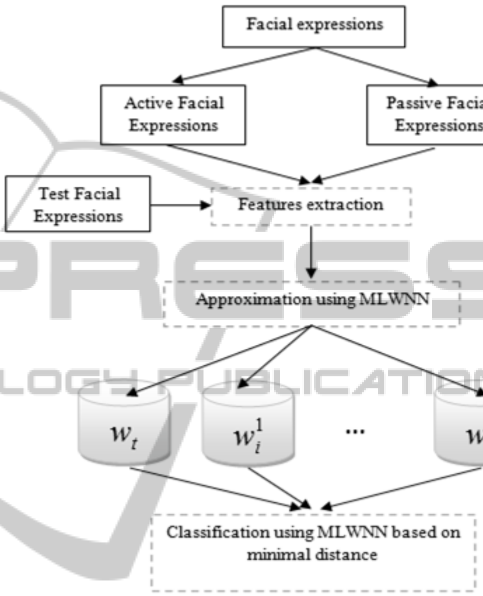

Figure 2: Proposed facial expression classification.

Figure 3: Manually landmark for different facial expressions.

4 FACIAL EXPRESSION REPRESENTATIONS

Facial expression representation is essentially a feature extraction process which converts representation of the face in terms of its landmarks. A landmark-based representation uses facial characteristic points which are located around specific facial areas, such as edges of eyes, nose, eyebrows and mouth as shown in figure 3. Since these areas show significant changes during facial articulation, we proposed a geometric face model based on 21 facial characteristic points for the frontal face view as shown in Figure 4 (a). Subsequently, thirteen characteristic distances are defined and estimated in Figure 4 (b). After facial feature extraction, a feature vector built from feature measurements, such as the brows distance, mouth height, mouth width etc., is created.

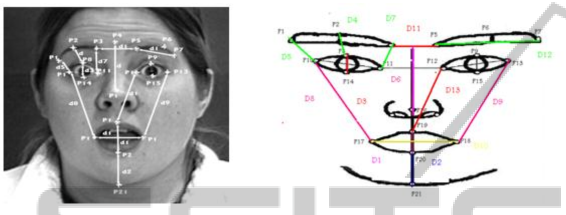
(a) (b)

Figure 4: Facial feature points and Geometric measures between facial feature points. D_i shows distances.

5 RESULTS AND DISCUSSION

To evaluate our method, we use Cohn Kanade expression database as input in our experiments. The database contains image sequences of over 97 subjects. 65% of them are females, while 35% are males. The motivation for the selection of this database originates from its content, i.e, the fact that image sequences depict the formation of human facial expressions from the neutral state to the fully expressive one.

First we considered to work with the same parameters as the Hammal method, i.e. with the thirteen characteristic distances shown in Figure 4 (b). To do this, we studied the variation of each distance comparing to the neutral face for each person of the database and for each emotion. An example of the results obtained for distance D1, D2,…, D13 is shown in Figure 5. From these distances we used Multi Library Wavelet Neural Network to classify emotions according to the value of the states (Table 1).

To demonstrate the suitability of our approach for facial expression recognition, we compare it with Hammal method (Marti, 2007). One can see that for Hammal approach anger, fear and sadness are not recognized but in the proposed approach only sadness is not recognized, because facial deformations between sadness and neutral are very small. For this reason we propose in future studies to add some other characteristics for sadness emotion to perform our approach.

Table 1: Classification rates of Hammal (second column) and of our method (third column).

EMOTION	% SUCCESS HAMMAL METHOD	% SUCCESS POPOSED METHOD
Joy	87.26	85.56
Surprise	84.44	97.93
Disgust	51.20	84.53
Anger	not recognized	89.69
Fear	not recognized	80.41
Sadness	not recognized	not recognized

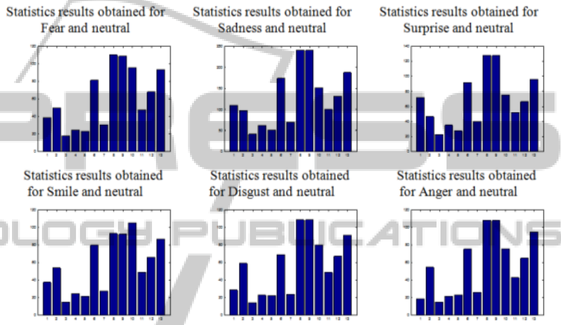

Figure 5: Statistics results for the 13 characteristic distances (x axis) obtained for the 6 emotions and neutral face.

6 CONCLUSIONS AND FUTURE WORK

We have presented a method based on wavelet neural network for automatic facial expressions classification. Simulation results show its simplicity and produce very significant improvement rates. The recognition rates for the proposed approach and Hammal one are respectively 85.56 versus 87.26 for joy, 97.93 against 84.44 for surprise and 84.53 over 51.20 for disgust. Anger and fear are recognized by the proposed approach; but both of them do not recognize sadness emotion. In the future, we hope to introduce new characteristics in the form of face distances or angles (for example the angle formed by the eyebrows). Another noticeable short-term objective is to track landmarks automatically using salient point techniques.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of this work by grants from General

Direction of Scientific Research (DGRST). Tunisia under the ARUB program.

REFERENCES

Cohen P., Johnston M., McGee D., Oviatt S., Clow J., Smith I., The efficiency of multimodal interaction: A case study. *in Proceedings of International Conference on Spoken Language Processing*. ICSLP'98. Australia. 1998.

Oviatt S., Ten myths of multimodal interaction. *Communications of the ACM. Volume 42*. Number 11 (1999). Pages 74-81.

Oviatt S., DeAngeli. A, Kuhn K., (1997) Integration and synchronization of input modes during multimodal human-computer interaction. *In Proceedings of Conf. Human Factors in Computing Systems CHI'97*. ACM Press. NY. pp. 415 – 422.

Jaimes A., and Sebe N., Multimodal Human Computer Interaction: A Survey. *IEEE International Workshop on Human Computer Interaction*. ICCV 2005. Beijing. China.

Mehrabian A., Communication without words. *Psychology Today*. vol. 2. pp. 53–56. 1968.

Zeng Z., Tu J., Liu M., Huang T. S., Pianfetti B., Roth D., Levinson S., Audio-Visual Affect Recognition. (2007) *IEEE Trans. Multimedia.* vol. 9. no. 2.

Pantic M., and Rothkrantz L. J. M., Towards an affect-sensitive multimodal human-computer interaction (2003). *Proc. of the IEEE*. vol. 91. no. 9. pp. 1370-1390.

Ioannou S., Raouzaiou A., Tzouvaras V., Mailis T., Karpouzis K., Kollias S., Emotion recognition through facial expression analysis based on a neurofuzzy network (2005). *Special Issue on Emotion: Understanding & Recognition Neural Networks*. Elsevier. Volume 18. Issue 4.pp. 423-435.

De Silva L. C., and Ng P. C., Bimodal emotion recognition (2000). *in Proc. Face and Gesture Recognition Conf.* pp. 332-335.

Gunes H., and Piccardi M., Fusing Face and Body Gesture for Machine Recognition of Emotions (2005). *IEEE International Workshop on Robots and Human Interactive Communication*. pp. 306 – 311.

Karpouzis K., Caridakis G., Kessous L., Amir N., Raouzaiou A., Malatesta L., Kollias S., Modeling naturalistic affective states via facial vocal and bodily expressions recognition (2007). *AI for Human Computing. LNAI* Volume 4451/2007. Springer.

Chihaoui M., Bellil W., and Amar C., (2010): Multimother wavelet neural network based on genetic algorithm for 1D and 2D functions approximation. *In Proceedings of the International Conference on Fuzzy Computation and 2nd International Conference on Neural Computation*. pp 429-434.

Bellil W., Othmani M., Ben Amar C., Alimi M. A., (2008) A new algorithm for initialization and training of beta multi-library wavelets neural network. *Aramburo J. Ramirez Trevino A Advances in robotics. automation and control*. *I-Tech Education and Publishing*. pp 199–220.

Bellil.W, Othmani. M, Ben Amar. C., Initialization by Selection for multi library wavelet neural network training (2007). *Proceeding of the 3rd international workshop on artificial neural networks and intelligent information processing ANNIIP07. in conjunction with 4th international conference on informatics in control. automation and robotics ICINCO*. pp.30–37.

Martí J., et al. (Eds.): IbPRIA (2007*). Part II. LNCS 4478*. pp. 40–47. 2007.© Springer-Verlag Berlin Heidelberg 2007.

Othmani M., Bellil W., Ben Amar C., and Alimi M. A., A novel approach for high dimension 3D object representation using Multi-Mother Wavelet Network (2011). *International Journal Multimedia Tools and Applications*, MTAP, Springer Netherlands, ISSN 1380-7501, pp. 1-18.