Cogn Neurodyn (2013) 7:47–57
DOI 10.1007/s11571-012-9216-y

RESEARCH ARTICLE

Analyzing the dynamics of emotional scene sequence using recurrent neuro-fuzzy network

Qing Zhang · Minho Lee

Received: 16 January 2012 / Revised: 13 July 2012 / Accepted: 6 August 2012 / Published online: 17 August 2012

Abstract In this paper, we propose a new framework to analyze the temporal dynamics of the emotional stimuli. For this framework, both electroencephalography signal and visual information are of great importance. The fusion of visual information with brain signals allows us to capture the users' emotional state. Thus we adopt previously proposed fuzzy-GIST as emotional feature to summarize the emotional feedback. In order to model the dynamics of the emotional stimuli sequence, we develop a recurrent neuro-fuzzy network for modeling the dynamic events of emotional dimensions including valence and arousal. It can incorporate human expertise by IF-THEN fuzzy rule while recurrent connections allow the fuzzy rules of network to see its own previous output. The results show that such a framework can interact with human subjects and generate arbitrary emotional sequences after learning the dynamics of an emotional sequence with enough number of samples.

Keywords Dynamics of emotion · Electroencephalography (EEG) · Fuzzy-GIST · International affective picture system (IAPS) · Recurrent neuro-fuzzy network (RNF)

Q. Zhang
China Samsung Telecom R&D Center, Chaoyang District, Beijing 100-125, China
e-mail: zhangqing.ee@gmail.com

M. Lee (✉)
School of Electrical Engineering and Computer Science, Kyungpook National University, 1370 Sankyuk-Dong, Puk-Gu, Taegu 702-701, South Korea
e-mail: mholee@knu.ac.kr

Introduction

The need for computational and robotic models which can understand the emotional state of the user is ever growing (Picard 2000). A lot of literatures are dedicated to studying affect detection. Calvo et al. (2010) give a comprehensive interdisciplinary review of models, methods and applications of affect detection. However, time is clearly important in emotion and emotions are a special dynamic form of cognition, many researchers start to be interested in the dynamical characteristics of emotions (D'Mello 2011; D'Mello and Graesser 2011). In this paper, we propose a new framework aiming to analyze the temporal dynamics of the emotional stimuli sequence.

In order to make the system be capable of understanding more complex emotions, we consider a valence-arousal (VA) model (Russell 1980; Lang 1995). By using such a dimensional approach, all emotions can be represented as points in the VA space, in which we can label the images joy, pleasure, anger and sadness. Not only is the VA space helpful in visualizing the location, extent and relationships between emotion categories, but also it is associated with the limbic system which suggestively supports a variety of functions including emotion, behavior, long term memory, and olfaction. Responses along the emotional valence dimension are associated with significant clusters in the amygdala, the anterior parietal cortex, and the insular cortex (Anders et al. 2004). Responses along the arousal dimension are associated with significant clusters in two regions: activity in the right supramarginal gyrus, and thalamic activity varied with reported arousal (Anders et al. 2004). For emotion related feature extraction, we adopted the fuzzy-GIST, which is a kind of conceptual gist of a scene that contains semantic information from both electroencephalography (EEG) and visual information

(Zhang and Lee 2012). Furthermore, we develop a novel recurrent neuro-fuzzy network to incorporate the human expertise to model the dynamic events of emotional stimuli sequence.

The remaining sections are organized as follows. "Methods" section introduces the proposed recurrent neuro-fuzzy network for this study including fuzzy-GIST. In "Experiments" section, we will give the experiment results and evaluate the performance of the proposed system. Some final conclusions and discussions are given in the last section.

Methods

Overview of the emotion dynamic analysis

Figure 1 demonstrates the graphic outline of the proposed approach. The considered input sequence is split into the sequences of visual information and EEG signal. The EEG can be allocated to specific image in the sequence, therefore, both EEG and visual features are dynamic (Gao et al. 2011). After signal processing, we extract the fuzzy-GIST for 2-emotion understanding to model the dynamics of emotional valence through a recurrent neuro-fuzzy network. Meanwhile, by taking arousal indicator into consideration, the fuzzy-GIST for 4-emotion understanding are fed to another recurrent neuro-fuzzy network to analyze the dynamic events of the emotional dimension of arousal. The fuzzy-GIST for 2-emotion and that for 4-emotion are introduced in Zhang and Lee (2012). We can then monitor the emotional trajectory by mapping each stimulus onto VA space.

Fuzzy-GIST as emotional feature at semantic level

Since we need features in emotional perspective, we propose the fuzzy-GIST to build a semantic feature vector to represent a scene image as well as consider the human feeling stimulated by the scene. The fuzzy-GIST is originated from the "GIST" (Oliva and Torralba 2006; Zhang and Lee 2009), and it is a kind of conceptual gist of a scene that contains semantic information. The procedure of extracting fuzzy-GIST from a natural scene is demonstrated in Fig. 2. 12-channel EEG signals are recorded and we adopt wavelet decomposition (WLD) (Burrus et al. 1997) for denoising EEG in our study. The selected wavelet filter for denoising the raw EEG signal is the reverse biothorgonal6.8 (rbio6.8) (Mallat 1989; Cong et al. 2012), and we select the D7, D8, and D9 to reconstruct the desired "real signals" (Zhang and Lee 2012). We focus on the 500 ms time course starting from stimulus onset and extract the power difference between left and right hemispheres in both alpha and gamma band to monitor the valence state of test subjects (Niemic 2002; Müller et al. 1999). On the other hand, beta/alpha ratio is used as an indicator if the subject is in an arousal state (Kandel et al. 2000). The EEG features from a subject are processed by the fuzzy C-means clustering (FCM). Based on the clustering result, a natural scene is assigned to positive/negative and calm/arousal groups to a degree of belongingness. According to the relation between the orientation distribution and human emotion evoked by a natural scene, the FCM is used to partition the orientation information of the image into 4 classes in terms of the orientation distribution and make an orientation descriptor for the image. We can describe the lightness as very dark, dark, middle, light and very light. A membership grade maps semantic words because of the fuzziness of human perception (Jang et al. 1997). In the similar way, we got warm-cool descriptor including warm, middle and cool, as well as saturation descriptor that indicates the low, middle and high saturation of the natural scene. The brain activity membership grades and visual information membership grades are cascaded to construct the emotional feature space. The difference between the fuzzy-GIST for 2-emotion

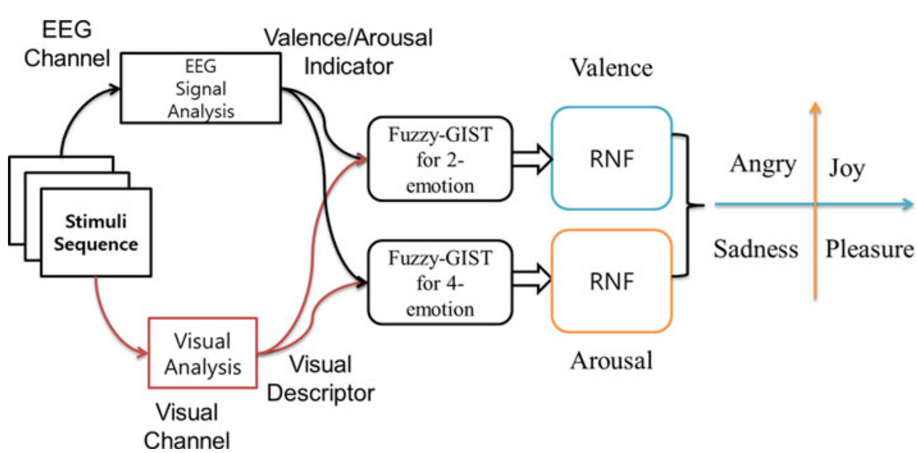


Fig. 1 Graphic outline of the proposed approach

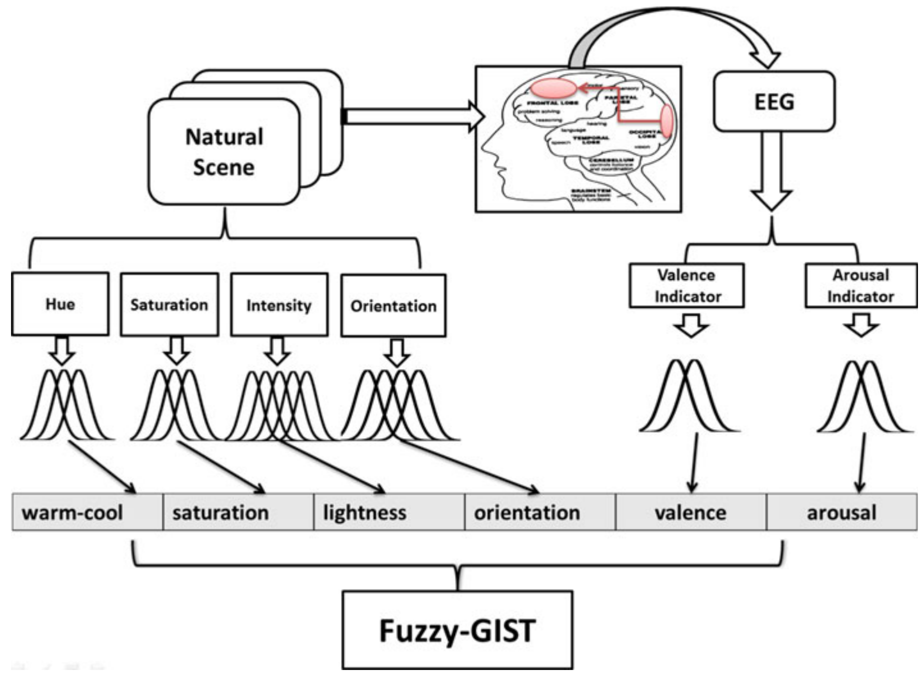


Fig. 2 The procedure of extracting fuzzy-GIST from a natural scene

understanding and that for 4-emotion understanding depends on considering the arousal indicator obtained from EEG signals. Different subjects may have different emotional responses even toward the same scene image, the fuzzy-GIST based on the combination of visual semantic information and the semantic EEG information can help to handle this personal bias for the emotion recognition.

Recurrent neuro-fuzzy network

The neuro-fuzzy network incorporates neural network learning concepts in fuzzy inference systems, resulting in adaptive neuro-fuzzy inference systems (ANFIS) (Jang 1993). In the ANFIS, the fuzzy system can be represented in a parametric form and the parameters are adaptively tuned by a learning procedure. ANFIS is a neuro-fuzzy network to recognize a pattern and adapt itself in a changing environment, while the fuzzy inference systems extract fuzzy rules and perform inferring and decision making (Held et al. 2006; Kecman 2001; Wang et al. 2005).

There is possibility, which is to allow time to be represented by the effect it has on processing. This means giving the processing system dynamic properties which are responsive to temporal sequences. In short, for modeling the dynamics of sequence, the network must be given temporal memory (Frayman and Wang 1998).

Figure 3 shows the proposed recurrent neuro-fuzzy network. As show in the Fig. 3, the proposed network is based on Takagi-Sugeno-Kang (TSK) type neuro-fuzzy inference system (Jang et al. 1997; Jang 1993; Nauck et al. 1997). The input of the network consists of previous and current input fuzzy-GIST, as well as previous output of the network (Gonzalez and Yang 2010; Wu et al. 2010). Nodes in layer 2 act as membership function to express the input fuzzy linguistic variables. The Gaussian membership function is adopted for the nodes in layer 2. Each node in layer 3 is called a rule node, it is formed by fuzzy AND operation. Nodes in layer 4 are called consequent nodes which perform a weighted linear combination of the input variables. The output of the network is the result of defuzzification of outputs of layer 4. The recurrent connections allow the network's fuzzy rule nodes to see their own previous output, so that the subsequent behavior can be shaped by previous responses. The recurrent connections are what give the networks memory.

To give a clear understanding of the mathematical function of each node, we will describe function of recurrent neuro-fuzzy network layer by layer. For notation convenience, the net input to the ith node in layer k is denoted by $u_i^{(k)}$ and the output value by $O_i^{(k)}$.

Layer 1: The node only transmits input values to layer 2. No function is performed in this layer.

$$O_i^{(1)} = u_i^{(1)} = [x(n), x(n-1), \ldots, x(n-p), y(n-1), y(n-2), \ldots, y(n-q)] \tag{1}$$

Layer 2: The following Gaussian membership function is used:

$$O_{ij}^{(2)} = exp\left\{ -\frac{\left(u_j^{(2)} - m_{ij}\right)^2}{\sigma_{ij}^2} \right\} \tag{2}$$

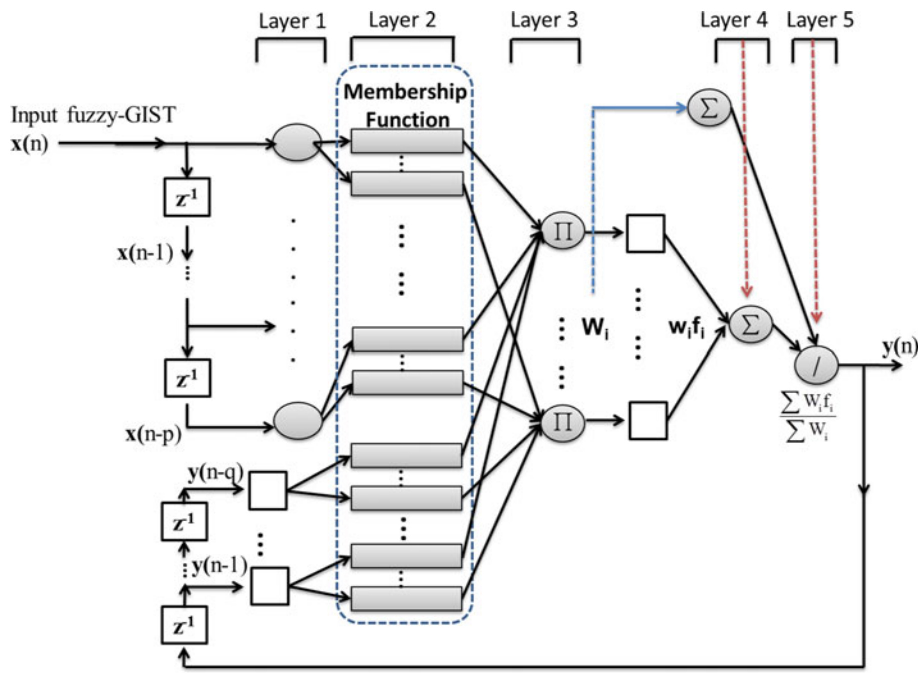


Fig. 3 The architecture of the proposed recurrent neuro-fuzzy network

where m_{ij} and σ_{ij} are the center and the width of the Gaussain membership function of the ith term of the jth input variable $u_j^{(2)}$.

Layer 3: The output of each node in this layer is determined by fuzzy AND operation. Here, the product operation is utilized to determine the firing strength of each rule. The each rule is defined as:

$$O_i^{(3)} = W_i = \prod_{j=1}^{N} O_{ij}^{(2)} = exp\left\{ -\sum_{j=1}^{N} \frac{\left(u_j^{(2)} - m_{ij}\right)^2}{\sigma_{ij}^2} \right\} \tag{3}$$

where N is the dimension of input layer.

Layer 4: The consequent node which performs a weighted linear combination of the input variables is defined as:

$$O_i^{(4)} = \sum W_i f_i = \sum W_i \left(b_{i0} + \sum_{j=0}^{p} a_{ij} x(n-j) + \sum_{k=1}^{q} a_{ik} y(n-k) \right) \tag{4}$$

where b_{i0} is a bias term.

Layer 5: The node in this layer computes the output signal y of the recurrent neuro-fuzzy network by defuzzification. The mathematical function is as:

$$y(n) = O^{(5)} = \frac{\sum W_i f_i}{\sum W_i} \tag{5}$$

The input consists of the fuzzy-GISTs of current stimulus and previous 2 stimuli, as well as outputs of previous 5 stimuli. As far as training is concerned, the backpropagation through time (BPTT) algorithm is used (Haykin 1998).

Experiments

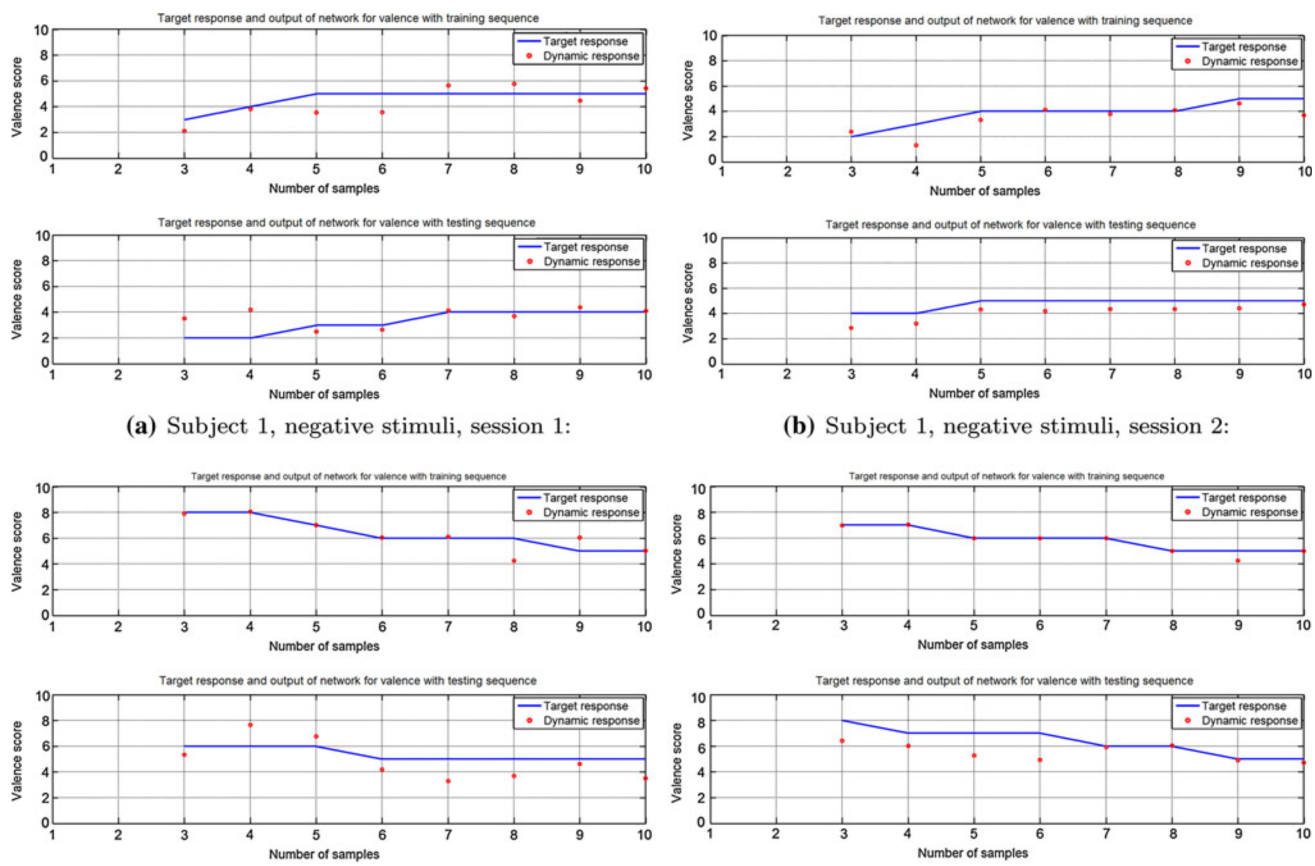
The dynamics of sequence consisting of identical/ similar emotional scenes

We first conducted a simple experiment to investigate the dynamics of emotional sequences that consist of same or similar emotional scenes. Five subjects participated in the experiments. We prepared 4 positive and 4 negative emotional sequences[1] for each subject, and each sequence consists of 10 identical emotional scenes.

Twelve channels of unipolar EEG (AF_Z, F_3, F_Z, F_4, FC_3, FC_Z, FC_4, T_7, C_3, C_Z, C_4, T_8) were obtained using EEG acquisition equipment BIOPAC MP150 and an electrode cap. Fpz was taken as the ground while the linked-earlobe played the role of the reference since this method used each ear as a reference for its own hemisphere and thus could reduce the electrocardiography (ECG) artifact. Impedance in each channel was kept below 10 kΩ. All channels were preprocessed on-line by a 0.1–100 Hz band pass filter and active notch filter to neglect the power line interference. And the data acquisition system was set to trigger at the appearance of visual

[1] For this experiment, we do not consider arousal axis for this experiment.

stimuli. The EEG with a sampling rate of 1,200 Hz and a total recording time of 4 s was used for observing the human brain activity stimulated by natural scene. During the whole experiment, we kept the environment as quiet as possible. The subjects were instructed not to blink, move eyes, or move any other part of the body but try to stay relaxed and keep the eyes open during the image appearance. In addition, before any recording for an image, we supplied a neutral background to help the participants be emotionless. The rbio6.8 wavelet filters were used to obtain the single trial Event Related Potentials (ERPs), and since the P100, N200 and P300 were obviously observed, we used the data which is a 500ms time course after the onset of the stimulus for the further feature extraction (Zhang and Lee 2012). The power difference between left and right hemispheres in both alpha and gamma bands were used to indicate the valence state of test subjects. In order to do this, we computed the power spectrum as for each channel to measure the power at various frequencies, and we used the average of left and right hemisphere power values at alpha (8–12 Hz) and gamma (30–40 Hz) bands to get the hemisphere power asymmetry. Then the valence indicators for a subject were processed by the FCM. Based on the clustering result, a natural scene is assigned to positive/negative categories to a degree of belongingness. The orientation information was obtained by a group of multi-scaled oriented filters. At different scale, they were of 4, 6 and 8 orientations, respectively. Then the basic orientation features were down-sampled to a size of 4×4 to form the orientation feature vector with dimension of $4 \times 4 \times 18$. The FCM is used to partition the orientation information of the image into 4 classes in term of the orientation distribution, and make an orientation descriptor for the image. The $L^*C^*H^*$ color space was used for color information, and they were down-sampled to a size of 10×10, respectively. Therefore, the 3 color feature vectors with dimension of 10×10 can be obtained. Each component (lightness, chroma and hue) was clustered into a specific number of groups to generate the semantic descriptors for the natural scenes. The lightness descriptor was used to abstract an image as very dark, dark, middle,



(a) Subject 1, negative stimuli, session 1:

(b) Subject 1, negative stimuli, session 2:

(c) Subject 1, positive stimuli, session 1:

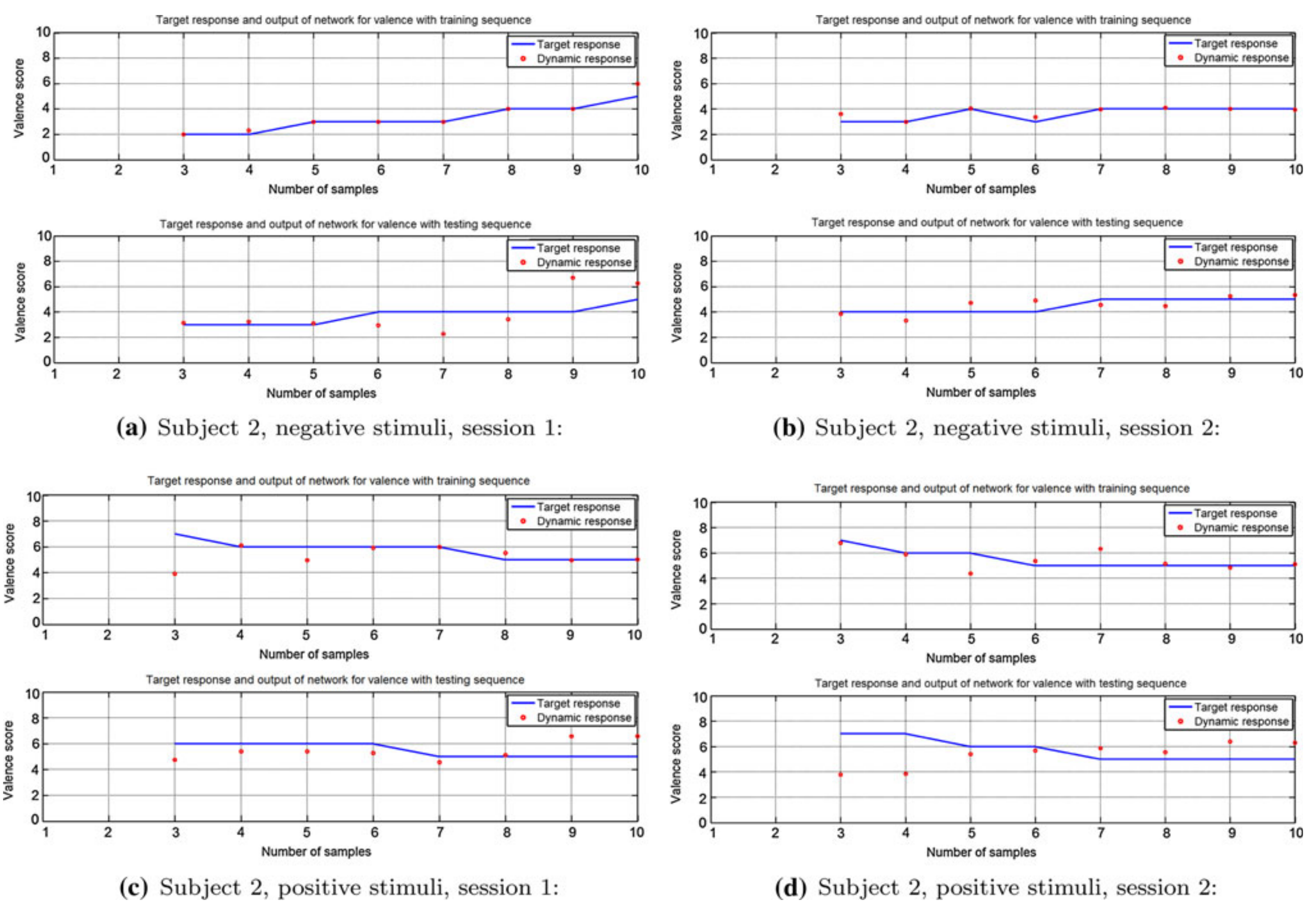
(d) Subject 1, positive stimuli, session 2:

Fig. 4 Modeling the dynamics of emotional sequences through interacting with Subject 1. **a** and **b** for negative sequences and **c**, **d** for positive sequences. For each sub-figure, the *upper* plot shows the training procedure for a emotional sequence, *blue line* represents target response and *red circles* represent response of dynamic neuro-fuzzy network; The *bottom* plot shows how we can generate dynamics for a new emotional sequence. (Color figure online) light and very light. In the similar way, we got saturation descriptor that indicates the low, middle and high saturation of a natural scene, as well as warm-cool descriptor indicating warm, middle and cool.

After signal processing, we extract the fuzzy-GIST for 2-emotion understanding to model the dynamics of emotional valence through a RNF network. Input of the RNF network consists of the fuzzy-GISTs of current stimulus and previous 2 stimuli, as well as outputs of previous 2 stimuli (e.g., $p = 2$, $q = 2$), which allows the model have enough dynamic input, but keep the dimension of input to model not too high so that the model will not have very sparse input. An emotional sequence is pair-wised with another according to emotional category. Therefore, after a RNF network learns dynamic events of an positive emotional sequence, it will be used to generate dynamics of another pair-wised positive sequence for testing. Using such an experimental paradigm, we aim to analyze the dynamic characteristics when an emotional stimulus or similar stimulus are repeatedly shown to subjects.

Figure 4a shows learning dynamics for a negative sequence and generating dynamics for a new negative sequence through interacting with subject 1. The blue solid lines represent target response while red dot circles represent response of dynamic neuro-fuzzy network. Figure 4b shows learning dynamics for another negative sequence and generating dynamics for a new negative sequence through interacting with subject 1. Figure 4c, d shows learning dynamics for positive sequences and generating dynamics for new positive sequences through interacting with subject 1. In the same way, analyses of dynamics for other three different subjects are shown in Figs. 5, 6 and 7.

As we can see from the results, when subjects repeatedly watch positive or negative stimuli, valence feedbacks from subjects vary and coverage at neutral state, this results in dynamics of emotion. In the same way, we observed the dynamics of emotional sequences for other four different subjects and the results were consistent (Results for three of them are shown above). Using the fuzzy-GIST and the RNF, we can successively model dynamic events of an emotional sequence.



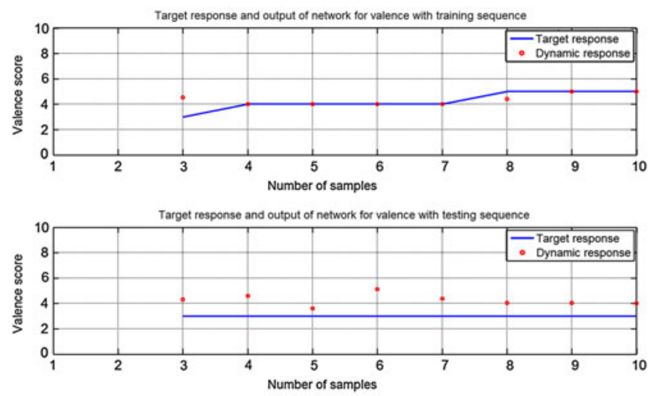
(a) Subject 2, negative stimuli, session 1:

(b) Subject 2, negative stimuli, session 2:

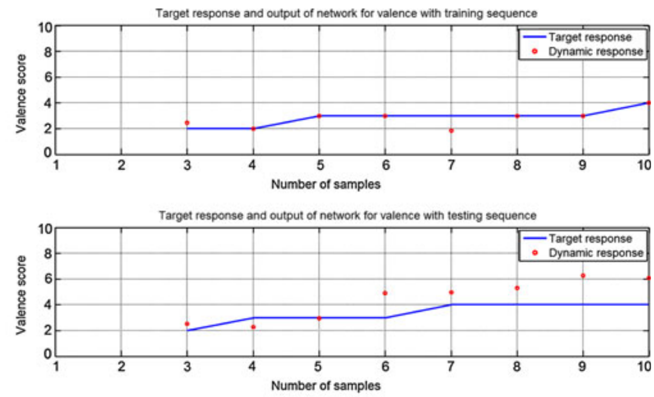
(c) Subject 2, positive stimuli, session 1:

(d) Subject 2, positive stimuli, session 2:

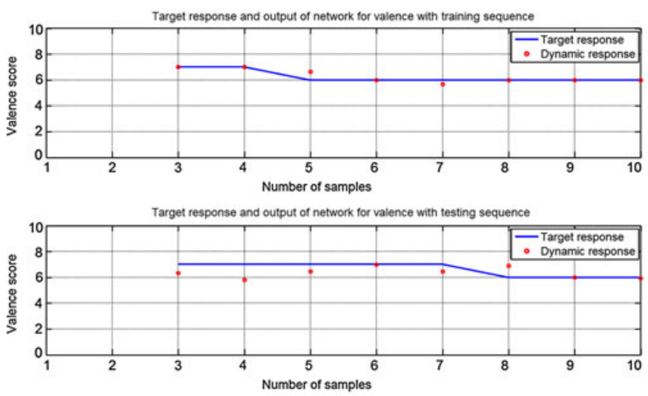
Fig. 5 Modeling the dynamics of emotional sequences through interacting with Subject 2. **a** and **b** for negative sequences and **c**, **d** for positive sequences. For each sub-figure, the *upper* plot shows the training procedure for a emotional sequence, *blue line* represents target response and *red circles* represent response of dynamic neuro-fuzzy network; The *bottom* plot shows how we can generate dynamics for a new emotional sequence. (Color figure online)



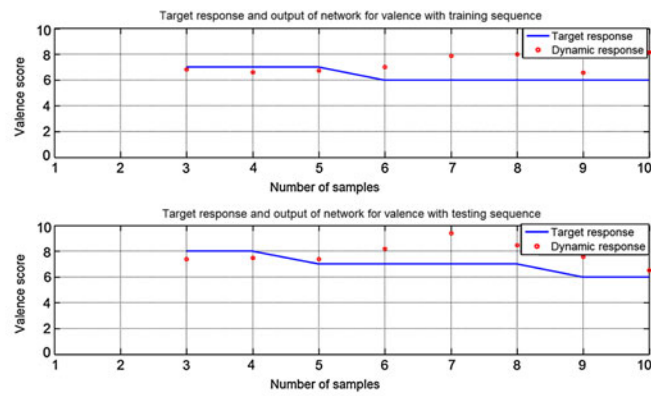
(a) Subject 3, negative stimuli, session 1:



(b) Subject 3, negative stimuli, session 2:



(c) Subject 3, positive stimuli, session 1:



(d) Subject 3, positive stimuli, session 2:

Fig. 6 Modeling the dynamics of emotional sequences through interacting with Subject 3. **a** and **b** for negative sequences and **c**, **d** for positive sequences. For each sub-figure, the *upper* plot shows the training procedure for a emotional sequence, *blue line* represents target response and *red circles* represent response of dynamic neuro-fuzzy network; The *bottom* plot shows how we can generate dynamics for a new emotional sequence. (Color figure online)

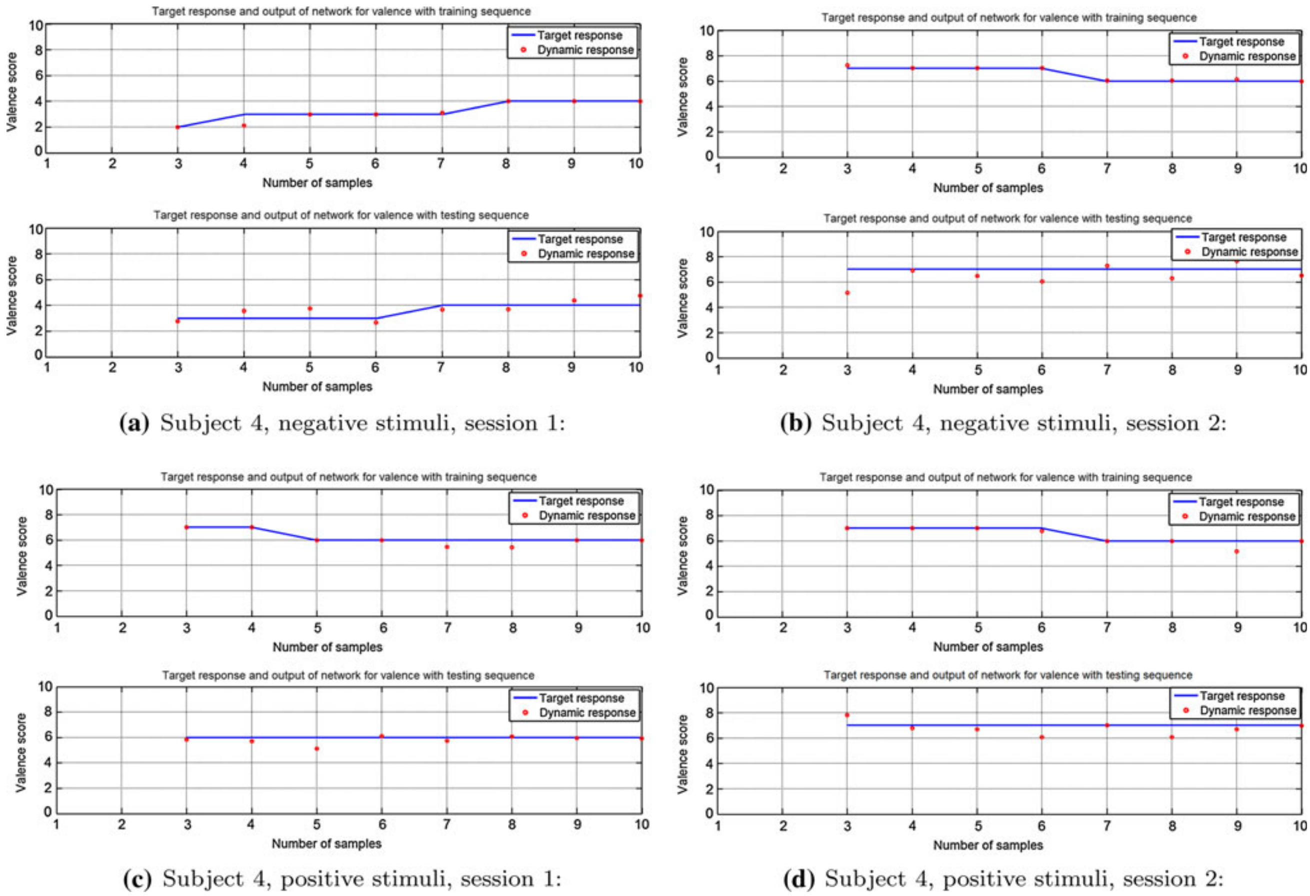
Analyzing the dynamics of complex emotional scene sequences

Next we did experiments to study the dynamics of more complex emotional sequences. 11 subjects participated in this study and 110 color images selected from IAPS and the corresponding EEG stimulated by images were used to extract the emotional features. In order to analyze the dynamic events of more complex emotional sequence, we adopt beta/alpha ratio as an indicator whether a subject is in arousal state. Because beta (13–30 Hz) waves are connected to an alert state of mind, whereas alpha waves are more dominant in a relaxed person (Kandel et al. 2000).

We randomly selected 50 % of the data and reordered them to generate 20 different emotional sequences for learning the valence and arousal dynamics of the emotional scene sequence. And we used the remaining data to generate random sequence to evaluate the generalization capability of the trained network. During the experiments, human subjects were asked to give valence and arousal scores to describe the emotional feedback of each stimulus. These results are used to evaluate the accuracy of an emotional sequence generated by the network. The current valence and arousal states highly depend on the previous states (Picard 1997). In order to handle the temporal effect on emotional processing, we took the average values of the valence and arousal scores of current stimulus and the previous two stimuli as valence value and arousal value for the current sample data. Thus, the valence and arousal scores become more smooth.

The prediction of emotional categories has two steps. In the first step, regression models were used to regress the valence/or arousal scores of an emotional stimulus. In the second step, emotional categorization was done by locating each stimulus in the VA space, according to the Algorithm 1.

Figure 8 shows the learning of dynamic events of an emotional sequence with interacting with a particular subject. The two subfigures on the top show the learning of valence and arousal dynamics, respectively. The solid lines represent the target responses while the dashed lines represent the dynamic responses of the RNF networks. On the bottom , the figure on the right-hand side is the emotional



(a) Subject 4, negative stimuli, session 1: **(b)** Subject 4, negative stimuli, session 2: **(c)** Subject 4, positive stimuli, session 1: **(d)** Subject 4, positive stimuli, session 2:

Fig. 7 Modeling the dynamics of emotional sequences through interacting with Subject 4. **a** and **b** for negative sequences and **c**, **d** for positive sequences. For each sub-figure, the *upper* plot shows the training procedure for a emotional sequence, *blue line* represents target response and *red circles* represent response of dynamic neuro-fuzzy network; The *bottom* plot shows how we can generate dynamics for a new emotional sequence. (Color figure online)

Algorithm 1 Algorithms for predicting the temporal emotional categories

Data: fuzzy-GIST for 2-emotion and 4-emotion;
Use RNF models to regress valence and arousal scores valence(n), arousal(n);
If *valence*(*n*) > = 5 **then**
  If *arousal*(*n*) > = 5 **then**
    emotion(n) = "Joy";
  else
    emotion(n) = "Pleasure";
  end
else
  If *arousal*(*n*) > = 5 **then**
    emotion(n) = "Angry";
  else
    emotion(n) = "Sadness";
  end
end

trajectory made by locating each stimulus in the VA space. Compared with the target emotional trajectory shown on the left-hand side, we can see that the networks is capable of learning the dynamics of an emotional sequence.

After training the RNF networks for valence and arousal dimensions, we used unlearnt data samples to randomly generate five emotional sequences with different lengths and different orders of emotional categories. Figures 9 and 10 demonstrate the results of the trained RNF networks generating arbitrary emotional sequences by interacting with the above mentioned subject.

As we can see from Figs. 9 and 10, through modeling the dynamics of valence and arousal dimensions separately, our proposed network is capable of generating the emotional trajectory for a new random emotion sequence.

We repeated the new sequence generation procedure five times for each subject, and then made a statistical result to evaluate the generalization performances of the framework by calculating the average accuracy for generating an emotional sequence as shown in Fig. 11. Here the accuracy indicates how many percent of emotional categories in a sequence can be correctly predicted.

As shown in the Fig. 11, we can see that the framework can interact with human subjects, learn the dynamic events of a sequence of emotional stimuli and then generate a new emotional sequence, which is close to the emotional feedback from human subjects. This study shows promising results revealing that machine is capable of interacting with human and generating its own emotion varying with natural scene stimuli.

Conclusion and discussion

The ultimate goal of this work is to build embedded emotion into a machine based on EEG and visual multimodal signals. First, we analyze emotion related information in the feature domain based on fuzzy-GIST, which considering the interaction between the brain signal and visual information. And also, the temporal dynamics of emotional sequences is exploited by a newly proposed RNF network, whose short term memory and approximation capabilities cater for modeling dynamic events in emotional sequences. This study shows promising results revealing that machine is capable of interacting with human and generating its own emotion varying with natural scene stimuli.

The proposed approach is likely to have a number of interesting applications such as intelligent tutoring system, human computer interface. It is an important property for human machine interaction, which tries to satisfy the affective users' requirements, and make them as productive as possible. Additionally, human emotion understanding system is critical because an affect sensitive interface can never respond to users' emotional states if it cannot understand their emotional states.

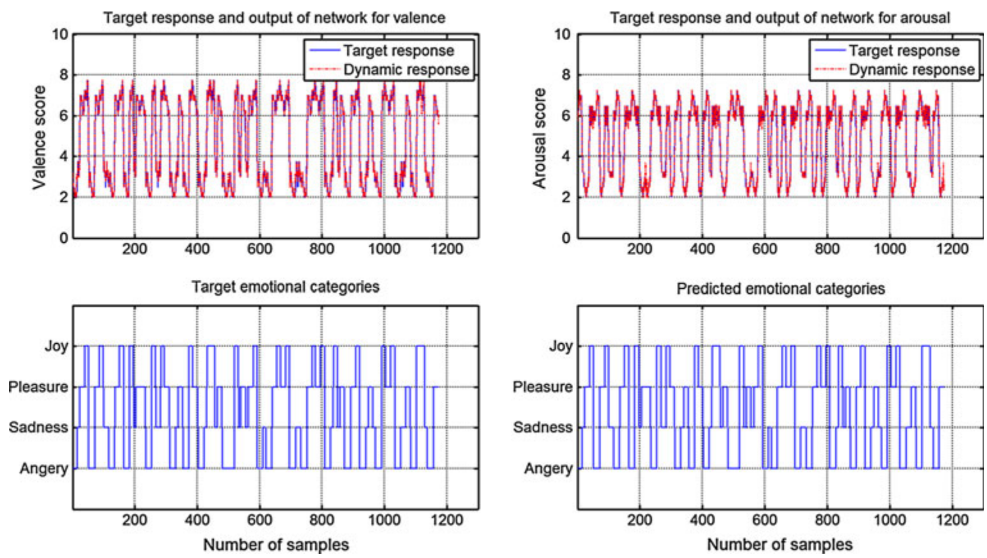


Fig. 8 Learning the dynamics of valence and arousal by two recurrent neuro-fuzzy networks

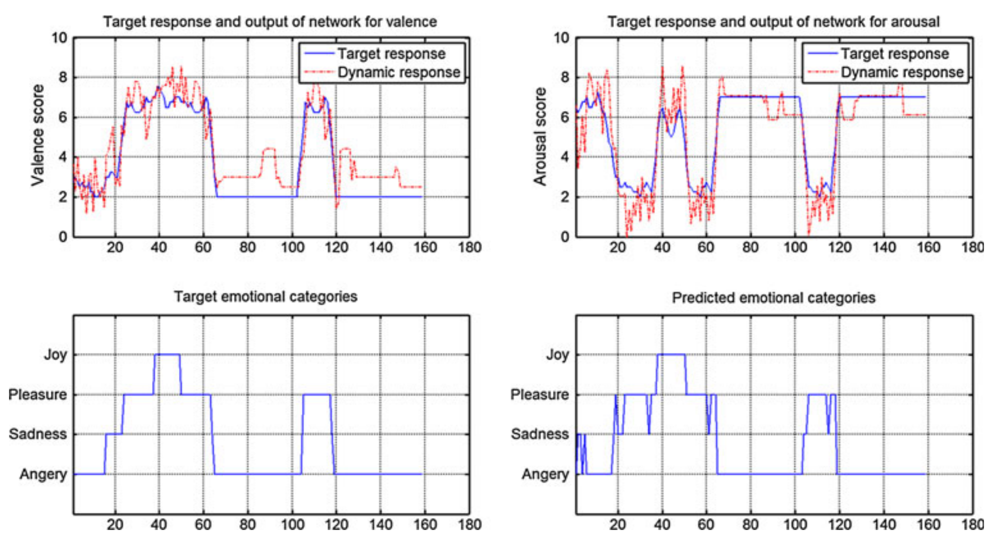


Fig. 9 Valence and arousal targets and output responses for a new sequence and its emotional trajectory

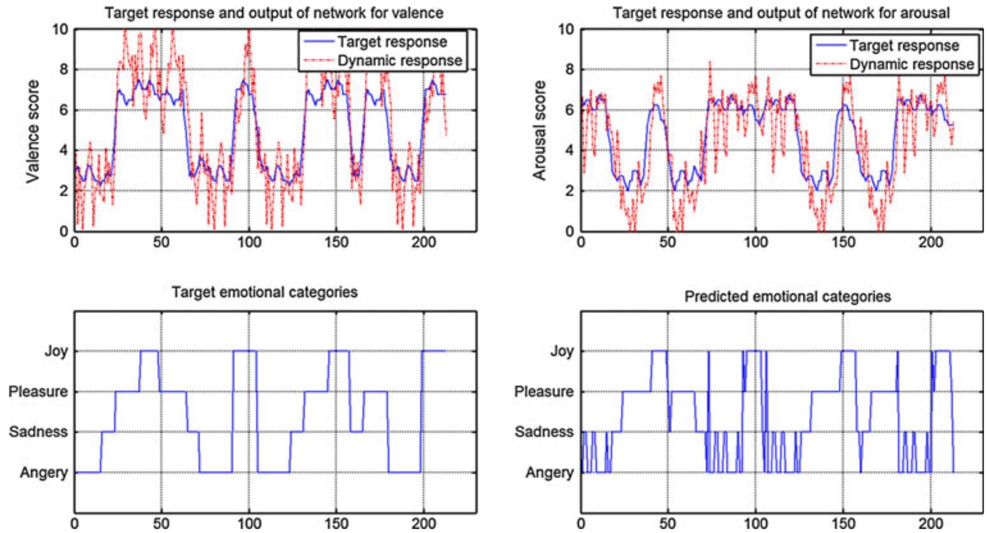


Fig. 10 Valence and arousal targets and output responses for another new sequence and its emotional trajectory

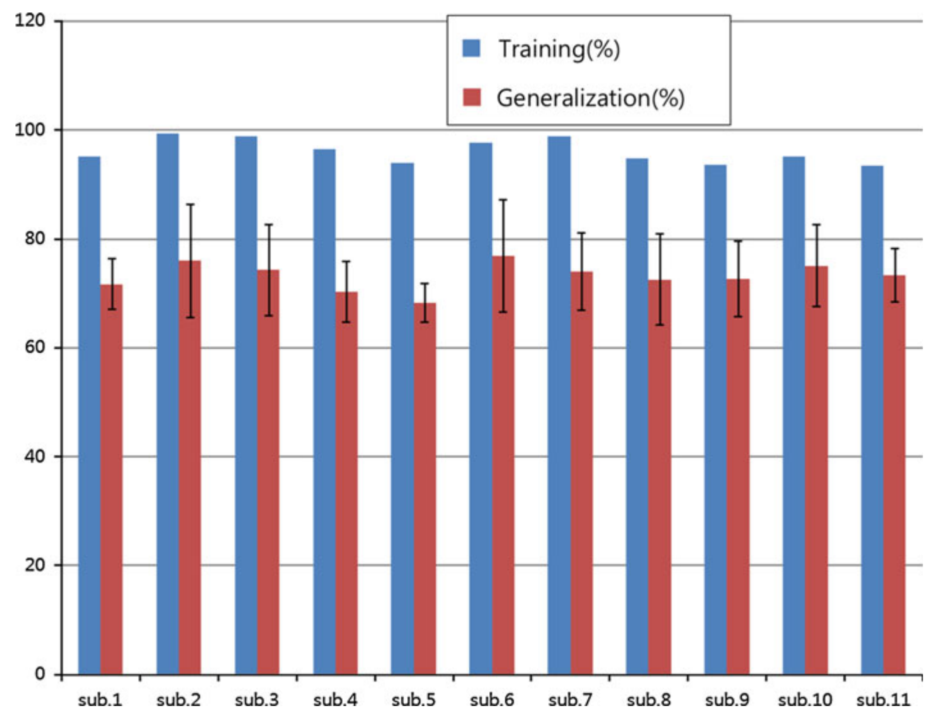


Fig. 11 Accuracy of generating an emotional sequence for training (%) and generalization [(mean ± SD)%]

Much work remains to be done in this largely unexplored field. First, we will more deeply investigate the development of emotional dynamics. In order to make a machine autonomously develop itself to learn the dynamics with new arriving data, an incremental learning algorithm for recurrent neuro-fuzzy network needs to be taken into consideration. Second, besides other kinds of emotional stimulus, audio cues also play an important role in emotion understanding for a machine. Therefore, we may include audio signal to form visual-audio stimuli for inducing subjects' emotional responses. In the next stage, we will try to consider the dynamic emotional feedback of a subject during watching a video clip. For this, we are going to develop a novel feature extraction method to extract 3-Dimensional emotional GIST for video clip. Third, we developed the current system interacting with human subjects through EEG signals. However, there exists many other modalities or channels for human machine interaction (e.g., face, voice, gesture, and other physiological signals). Each modality has advantages and disadvantages toward its use as a viable affect detection channel. In our future work, we may consider more different modalities. For example, besides using the EEG signals, we can also develop system to understand subjects' emotion state through their voice, body language, Electrocardiogram (ECG) signals, skin conductivity, etc.

Acknowledgments This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) (2012R1A1B6004217) (70 %), and by the Converging Research Center Program funded by the Ministry of Education, Science and Technology (2011K000659) (30 %).

References

Anders S, Lotze M, Erb M, Grodd W, Birbaumer N (2004) Brain activity underlying emotional valence and arousal: a response-related fMRI study. Hum Brain Mapp 23(4):200–209

Burrus CS, Gopinath RA, Guo H (1997) Introduction to wavelets and wavelets transforms. Prentice Hall, NJ, USA

Calvo RA, D'Mello S (2010) Affect detection: an interdisciplinary review of models, methods, and their applications. IEEE Trans Affect Comput 1(1):18–37

Cong F, Kalyakin I, Li H, Huttunen-Scott T, Huang Y, Lyytinen H, Ristaniemi T (2012) Answering six questions in extracting children's mismatch negativity through combining wavelet decomposition and independent component analysis. Cogn Neurodyn 5(4):343–359

D'Mello SK, Graesser AC (2011) Emotions during learning with AutoTutor. In: Durlach P, Lesgold A (eds) Adaptive technologies for training and education. Cambridge University Press, Cambridge (in press)

D'Mello SK (2011) Affect trajectories during complex learning. In: Seel N et al (eds) Encyclopedia of the sciences of learning. Springer, Berlin (in press)

Frayman Y, Wang LP (1998) Data mining using dynamically constructed recurrent fuzzy neural networks. Res Dev Knowl Discov Data Min 1394:122–131

Gonzalez-Olvera M, Tang Y (2010) Black-box identification of a class of nonlinear systems by a recurrent neurofuzzy network. IEEE Trans Neural Netw 21(4):672–679

Gao JB, Hu J, Tung WW (2011) Complexity measures of brain wave dynamics. Cogn Neurodyn 5:171–182

Haykin S (1998) Neural networks: a comprehensive foundation, 2nd edn. Prentice Hall, NJ, USA

Held CM, Heiss JE, Estevez PA, Perez CA, Garrido M, Algarin C, Peirano P (2006) Extracting fuzzy rules from polysomnographic recordings for infant sleep classification. IEEE Trans Biomed Eng 53:1954–1962

Jang J-SR, Sun C-T, Mizutani E (1997) Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice Hall, Inc., Upper Saddle River

Jang J-SR (1993) ANFIS: adaptive-network-based fuzzy inference systems. IEEE Trans Syst Man Cybern 23:665–685

Kandel ER, Schwartz JH, Jessell TM (2000) Principles of neural science. McGraw-Hill Medical, New York

Kecman V (2001) Learning and soft computing, support vector machines, neural networks and fuzzy logic models. The MIT Press, Cambridge

Lang PJ (1995) The emotion probe. Studies of motivation and attention. Am Psychol 50(5):372–385

Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693

Müller MM, Keil A, Gruber T, Elbert T (1999) Processing of affective pictures modulates right-hemispheric gamma band EEG activity. Clin Neurophysiol 110(11):1913–1920

Nauck D, Klawonn F, Kruse R (1997) Foundations of neuro-fuzzy systems. Wiley, Chichester

Niemic CP (2002) A theoretical and empirical review of psycho-physiological studies of emotion. J Undergrad Res 1:15–18

Oliva A, Torralba A (2006) Building the gist of a scene: the role of global image features in recognition. Prog Brain Res 155:22–36

Picard R (2000) Toward computers that recognize and respond to user emotion. IBM Syst J 39(3–4):705–719

Picard R (1997) Affective computing. MIT Press, London

Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39:1161–1178

Wu X, Wang Y, Huang L, Zuo Y (2010) Robust stability analysis of delayed Takagi-Sugeno fuzzy Hopfield neural networks with discontinuous activation functions. Cogn Neurodyn 4(4): 347–354

Wang LP, Fu XJ (2005) Data mining with computational intelligence. Springer, Berlin

Zhang Q, Lee M (2009) Analysis of positive and negative emotions in natural scene using brain activity and gist. Neurocomputing 72(4–6):1302–1306

Zhang Q, Lee M (2012) Emotion development system by interacting with human eeg and natural scene understanding. Cogn Syst Res 14(1):37–49