Preprint Version. Manuscript submitted to and accepted by 2016 IEEE International Conference on Systems, Man, and Cybernetics.

Assessing Haptic Video Interaction with Neurocognitive Tools

Shahzad Rasool*, Xiyuan Hou*, Yisi Liu*, Alexei Sourin**, Olga Sourina*

* Nanyang Technological University/FraunhoferIDM@NTU, Singapore, Singapore

** School of Computer Science and Engineering, Nanyang Technological University, Singapore

{shahzadrasool | HouXY | LiuYS | assourin | eosourina}@ntu.edu.sg

***Abstract*—Haptic interaction is a form of a user-computer interaction where physical forces are delivered to the user via vibrations, displacements and rotations of special haptic devices. When quality of the experience of the haptic interaction is assessed, mostly subjective tests using various questionnaires are performed. We proposed novel neurocognitive tools for assessing both overall experience of the haptic interaction, as well as particular time-stamped activities. Our assessment tools are based on recognition of emotions and stress obtained from Electroencephalograms (EEG). We used them in a feasibility study on adding haptic interaction to Skype video conversation.**

Keywords—EEG; neuroscience; user study; haptic interaction

I. Introduction

In computing, haptic interaction, or haptics, is implemented by using various haptic devices delivering forces to the users through vibrations as well as displacements and rotations of haptic actuators. Most frequently, desktop haptic devices and game controllers, as well as wearable vibrating mini-motors are used.

As with any other human-computer interaction, to evaluate how believable the haptic interaction is, subjective or objective Quality of Experience (QoE) user tests are normally done. Subjective user tests are performed by filling in various questionnaires to evaluate the impact of the overall haptic feedback. These questionnaires normally evaluate four factors: sensory immersion, comfort, realism, and satisfaction. Objective tests are not common and they measure changes in physiological state of body such as heart rate, respiration rate, body temperature, etc.

To improve the reliability and quality of the haptic experience tests, we propose novel neurocognitive tools based on recognition of emotions and stress obtained from electroencephalograms (EEG). These EEG-based methods can perform measurements of the user performance and satisfaction during the haptic interaction at any time interval starting from 1/32 sec thus giving objective and time-stamped feedback of the whole process of interaction. In contrast to post-experiment questionnaire-based evaluations, the EEG-based tools will not only evaluate the overall experience but also show us at which part of the interaction the users have become excited or bored, where exactly the mental load has got increased or positive emotions have been created. We hypothesize that with this approach we will not only evaluate the developed software but will also be able to actively engage the potential users into the design and development phases of various haptic interactions.

We collected EEG data in a feasibility study on adding haptic interaction component to Skype video conversation over the internet. We then processed these data and compared the obtained results with the results obtained from traditional questionnaires. The hypotheses that EEG-based methods can be used as objective way of quality assurance of haptic interaction was validated.

II. Related Works

A. *Subjective and Objective Methods in Haptics*

Common techniques of measuring subjective QoE in haptic multimedia applications are based on questionnaires. Objective measurements on the other hand obtain physiological data of the user to further estimate psychological and emotional states.

Questionnaires are usually the mode of choice in subjective evaluation of perception or presence in haptic systems. Typically, participants are asked to rank their experiences on Likert-type scales. For example, Rehman et al. [1] measured user satisfaction by requesting responses to seven questions on a 7-point Likert scale. These questions demanded feedback regarding acceptance of the system, comfort of usage, user interface, trainability, etc. Kim et al. [2] evaluated with questionnaires the effects of enhancing the movie experience through vibrotactile feedback. Multiple factors such as fatigue, usefulness, intuitivism, realism were evaluated as a measure of QoE using questionnaires. Different questions have been used to acquire user feedback through questionnaires however, they are mostly used to assess the four factors identified by Danieau et al. [3] namely: ‘Comfort’ for measuring acceptance of a system, ‘Satisfaction’ for estimating how enjoyable was the system, ‘Realism’ for assessing the consistency of the haptic feedback with the experience in real world, and ‘Sensory’ for describing how much the haptic feedback improved the immersion.

Methods of objective evaluation of haptic perception measure the physiological changes which the user experiences while using a haptic system. For example, heart rate and skin temperature were used by Meehan et al. [4] to infer presence in a virtual environment. Similarly, a number of physiological features such as heart rate, muscle activity, respiration rate, etc. were captured for video game players to estimate user experience. The measured change in physiological state does not directly reflect QoE but is used to determine the changes in emotional state of the user. These studies usually chose similar features for physiological measurement such as heart rate, skin conductance, facial expression, etc. in different virtual environments. In contrast to the existing approaches, we propose using neurocognitive tools to estimate emotional or psychological state of the users. These tools are based on recognizing emotion and stress from EEG signals.

B. EEG-based Methods and Their Current Uses

EEG, which measures the brain activity of humans, is increasingly popular nowadays. The current EEG devices on the market are affordable and comfortable to wear (being also wireless), which opens new application areas where brain states need to be recognized.

Emotions and stress are commonly recognized from EEG-based brain state. For emotions, dimensional and discrete ways for defining the feelings are used. In the dimensional way, an emotion is partitioned into three dimensions, namely arousal (ranging from calm to excited), valence (ranging from negative to positive), and dominance (ranging from low in control to full control). Alternatively, an emotion can be directly defined by given a label such as happy, sad, etc. In either case, the EEG-based emotion recognition proves its reliability. In [5], it is shown that by extracting power features from 32 channels of EEG data and using Support Vector Machine (SVM) as a classifier, four emotions can be recognized with 82.37% accuracy. In [6], where power feature and SVM were also used, the accuracy of 62.07% was achieved for three emotions recognition while the channel number was reduced to 16 channels. However, in both works offline processing with relatively large amount of channels is used while the target of EEG-based emotion recognition is to do it in real time and with as few channels as possible. In our previous works [7-9], subject-dependent emotion recognition algorithms were proposed. We were able to recognize positive and negative emotions in real time with only five channels, and the best recognition accuracy was 92.03% [9]. Higuchi fractal dimension [10], four band power features (theta, alpha, beta, theta/beta ratio), statistical features [11], and higher order crossings [12] were extracted as features and SVM was used as a classifier. It required a calibration to be done for each subject before the recognition.

Stress refers to the feeling of strain and pressure. It can also be recognized from EEG signals. In [13], higher order spectra and SVM were used to identify two stress level. In [14], the Higuchi's fractal dimension of EEG, Gaussian mixtures of EEG spectrogram, and Magnitude Square Coherence Estimation (MSCE) between the EEG channels were used as features and SVM as well as k-NN were used as classifiers. In our previous work [15], by using the combination of fractal dimension and statistical features as well as by applying the SVM classifier, four levels of stress were recognized with the average accuracy of 67.07%.

Furthermore, in work [17], we analyzed relation between mental workload calculated using traditional NASA-TLX method and an EEG-based recognition algorithm and found out that the data were highly correlated in most of the simulations implemented.

C. Research hypothesis on using EEG-based Method in Haptics Quality Experience Evaluation

Based on the existing ways of using EEG for identification of the brain states, we hypothesized that we should be able to capture changes of the users' feelings during haptic interaction. Common user questionaries evaluating haptic interaction can be used as a reference to be interpreted as emotion and stress level. Using time-stamped EEG data, we will also be able to actively engage the potential users into the design and development phases of various haptic interactions.

To prove the hypothesis, we applied EEG-based methods for evaluating a feasibility study on adding haptic interaction modality to common audio-video conversation with Skype. An experimental protocol has been developed to assess various visual and haptic experiences during such conversation and tangible interaction over the Internet.

III. EXPERIMENT SETUP

A. System Outline

We used two networked computers placed in to two different rooms. Each computer had a Geomagic Touch device (a desktop robotic hand used for navigation of 3D cursor and delivering forces back to the user) as well as a web camera, a microphone and headphones (Fig. 1).

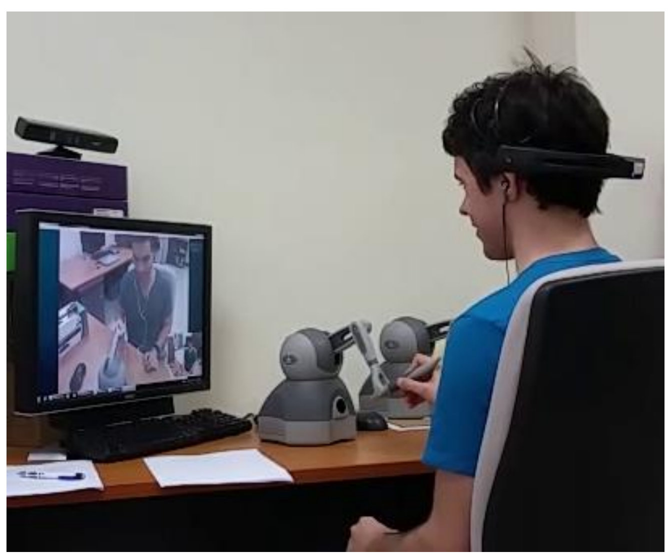


Figure 1. A participant of the experiment engaged in a video conversation while communicating motions across the network using a haptic device.

A TCP/IP connection was established between the two computers and the haptic device connected to each of them was programmed to follow the other party hand motions. This was done by transmitting the haptic device coordinates from one computer to another while both computers used the same coordinate system. Based on the proximity of the device coordinates received over the network to the coordinates retrieved from the remote haptic device, the remote computer calculated a magnetic force that attracted the haptic interface points of the two devices towards each other. In other words, a force field existed around the position of each device in the coordinate space used. As the two force fields overlapped, the remote computer applied the force on the remote device and transmitted the inverted force vector to the local computer. As a result, motion of a device handle on one computer produced the same motion of the remote device handle. A Skype video connection was also established between the two computers so that the user could see how their hands as well as haptic device handles move. Each participant wore an EEG headset during the experiment as shown in Fig. 1.

B. Participants

We used 10 participants, 9 males and 1 female, aged from 24 to 55 (M=31.4 SD=8.9). All participants knew each other. Two were experienced haptic users while others had little or no such experience. In order to diminish the 'surprise effect', all novice users were given 5 minutes of induction haptic exercises. The participants were grouped into 5 pairs based on their cultural, educational and language similarities to avoid possible emotional tensions that can affect the experiment.

C. Calibration of EEG-based Recognition

As both emotion and stress recognition are subject-dependent algorithms, two calibrations were carried out to train the classifiers. The targeted emotions to be recognized, while using haptics, are solely linked to different levels of valence (ranging from positive to neutral and neutral to negative). Thus we selected sound clips from IADS database [16] according to their valence ratings to evoke 7 different levels of valence. We carried out 7 sessions for emotion calibration and each lasted for 52 seconds (a 16 seconds silence followed by 6 sound clips with 6 seconds per clip).

For stress calibration, the Stroop color-word test was used. Four sessions were conducted to recognize four different levels of stress: absence of stress, low, moderate, and high stress. In the first session, the subject had to be simply relaxed, i.e. showing absence of stress. In the second session, the subject was asked to click the numeric key that represented the color of the word shown on the screen (the meaning of the word is the same with the color) which was supposed to elicit low stress. In the third session, the subject had to click the numeric key that represented the color of the word shown on the screen while the meaning of the word was different from the color (moderate stress). In the third session, the subject repeated the previous test but within a limited time (high stress).

The EEG signals of the two calibrations were recorded from the 14 channels of the Emotiv EPOC device with a sampling frequency of 128Hz and 16 bit A/D resolution. The Emotiv EPOC device used is convenient to set up for recording, and it provides comparable performance to a conventional EEG device at an affordable price [18]. The 14 electrode positions used are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, according to the 10-20 international system. During the recording, the raw EEG signals were transmitted wirelesly to the computer. Based on our previous work on emotion [7, 8] and stress recognition [15], we extracted the Higuchi fractal dimension [10], band power features (theta, alpha, beta), and statistical features [11] from the EEG data with a sliding window of size of 4 seconds and overlapping rate of 75%. Then, the features were fed into the SVM classifiers to train models. The models were saved and used to identify the brain states of the users while they were engaged in haptic interaction.

D. Procedure

After familiarization with the haptic devices followed by EEG calibration, each pair of the testers was asked to perform several activities in a single session which was divided into four 5-minute intervals. EEG data was recorded for each of the 5 minute intervals. The whole experiment lasted 30-45 minutes per pair of the participants.

1. The participants were asked to have a normal Skype audio-video conversation, without any haptic feedback. They were free to choose the topic and language of their choice.

2. While continue to be in a video conversation, both participants were asked to use the haptic device to exchange different hand motions representing handshake, holding hands together, vibrations, etc.

3. One user was asked to lead the conversation by being an active haptic user while the other user was asked to be a reactive user. The participants were given examples of activities that they could perform such as: /1/ to ask the reactive user to place his/her hand or finger in a certain position while trying to touch it with the remote device, /2/ to ask the reactive user to recognize the shape or letter remotely drawn by the active user with the device stylus. They could also come up with other interaction activities while maintaining active-reactive user pattern.

4. The same as in task 3 but the reversed active-reactive roles.

E. Subjective Measurements

After each activity, the participants were asked to fill up a

TABLE I.
SUBJECTIVE EVALUATION QUESTIONNAIRE. QUESTIONS ARE RATED ON A FIVE POINT LIKERT SCALE (1-NOT AT ALL, 5-TOTALLY).

No.	Factor	Question
1	Comfort	How much was the system comfortable?
2	Satisfaction	How much was the system pleasant to use?
3	Sensory	How much did haptic feedback improve the interaction?
4	Realism	How much did your experiences in the virtual environment seem consistent with your real-world experiences?

questionnaire which was designed to assess the four factors discussed in Section II, i.e. Comfort, Satisfaction, Realism and Sensory [3]. We considered Comfort, Realism and Sensory as measures of 'Presence', which is a suitable factor to measure the presented scenario of video conversation with haptic feedback. Satisfaction is a measure of usability. Other factors that estimate usability, such as efficiency and effectiveness, were not considered, as the experiment was rather open and not designed for performing only a specific task. The questions given in Table 1 were rated on a 5-point Likert scale.

The questionnaire for the first interval did not include Question 3 and 4 as no haptic feedback was provided. An optional question requesting comments was included to each questionnaire in the words "Any other comments about what you liked, did not like, or things that should be changed?" This was included to get participant feedback for improving the design of the system.

F. Objective Measurement

The EEG signals of the participants were recorded by Emotiv device while they were performing the activities in the sessions. Then, the same features were extracted as in the calibration, and the trained SVM models were applied to identify the emotional and stress states of the subjects during the session.

The obtained brain states were analyzed from two perspectives: /1/ an overall emotional state and stress level over each 5-minute activity; /2/ detailed emotional states and stress levels changes over every 10 seconds in each 5-minute activity. The first analysis is to be used to study how brain states change from one activity to another. The second analysis is to be used to study how brain states change within each activity.

IV. Analysis of Results

The data collected through questionnaires with respect to the four factors (sensory, realism, comfort and satisfaction) has shown that the addition of haptic interaction modality to video conversation improves satisfaction. In Fig. 2, the normalized mean opinion scores given by participants are shown for each interval. Although the level of satisfaction is slightly increased with the addition of haptic modality, the reported comfort level does not vary significantly across different intervals.

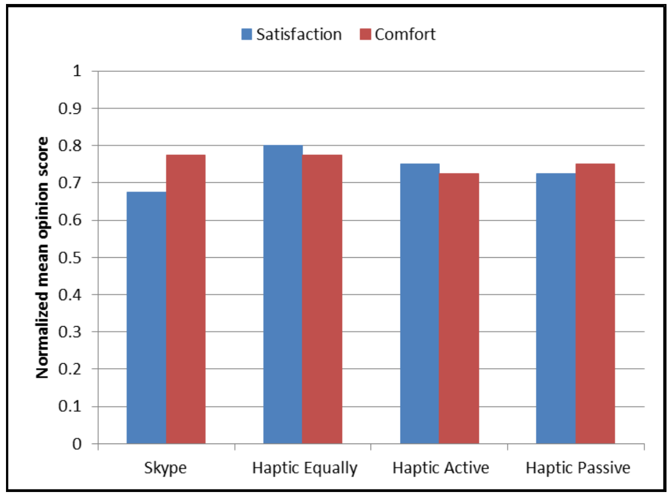


Figure 2. Normalized mean opinion score of all participants for the experiment with and without the haptic feedback.

Next, we analyzed the questionnaire responses to each individual interval. Fig. 3 shows the normalized opinion score for the three intervals (Equal, active and passive participations) with haptic feedback. The change in mean opinion score for realism and sensory factors is insignificant across different intervals, hence they are different activities.

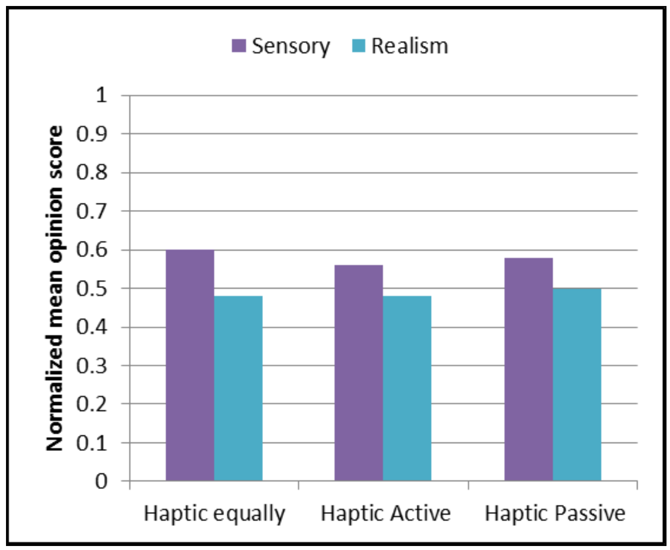


Figure 3. Normalized mean opinion score of all participants for different intervals of the experiment.

For objective measurement, the EEG signals of all participants were processed to recognize emotional states and stress levels in two ways: the first method gives an overall averaged emotional state and stress level over each 5-minute trial (Fig. 4) and the second method gives the averaged by 10 seconds values. In Fig. 4, the normalized mean values of valence and stress values over each 5-minute interval are given. It shows that the participants reach higher average valence level (more positive) when they were leading the conversion as the "active" haptic users. For the remaining three conditions, the valence levels of emotions are similar. As for the stress, it was slightly higher at the beginning of the experiment session during conversation without haptic interaction. During the next section, it was slightly reduced despite that both participants were equally involved in haptic interaction and conversation. The lowest stress level participants had when they were active haptic participants, that is consistent with the finding that they also experienced more positive emotions in this activity. As it follows from the EEG-based emotion and stress recognition results, the additional active haptic interaction could invoke more positive emotions and release stress during the audio-video conversation. In Fig. 2 and Fig. 4, the patterns of satisfaction (from Questionnaire) and valence (from EEG) for haptic active and haptic passive are similar. Most of the participants feel more satisfied and more positive during the active haptic interaction then during the passive haptic interaction.

The advantage of EEG-based brain states recognition is to provide detailed and quantified measurements during the experiment. Thus, in the second analysis, detailed emotional states and stress levels changes over every 10 seconds in each 5-minute trial were calculated. Fig. 5 describes a complete interaction process between two participants (subjects 5 and 6) using continuous changes of emotions (Fig. 5a) and stress levels (Fig. 5b) averaged by 10 seconds. Four intervals corresponding to 4 sections (each lasts about 300 seconds) are split by vertical lines. The first two intervals are traditional Skype conversation (0-300 seconds) and equal haptic Skype conversation (300-600 seconds). Subject 5 leads the conversation as active haptic user in the third interval (600-900 seconds), and then Subject 6 turns to be an active haptic user in the fourth interval (900-1200 seconds).

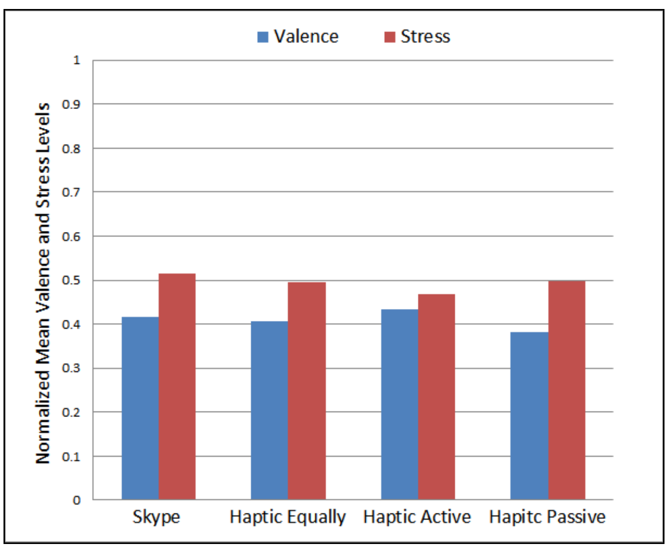
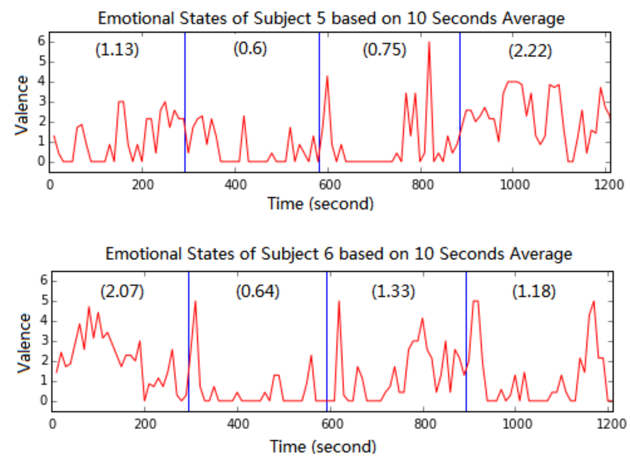


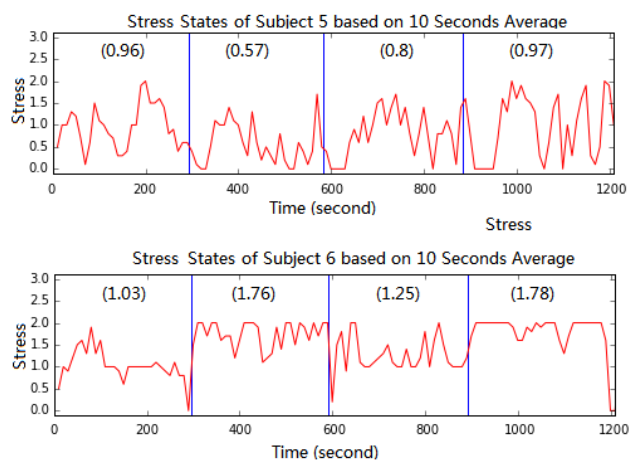
Figure 4. Mean valence and stress levels of all participants in four intervals of experiments

We then interviewed these two participants after the experiment. They were experienced haptic users. They commented that the second interval of the conversation with equal haptic interaction was a little boring for them, since they already did it before. However, in the first half of third interval, Subject 5 reached very low valence level from 600 to 750 seconds since he noticed a problem in the haptic interaction—navigation of the remote device while looking in the video image was not working like looking in a mirror, i.e. when moving the handle to the right, it moved to the left in the video, and visa versa. His stress level had also raised during this period of time. After a short conversation with Subject 6, Subject 5 had found a solution around the 800 second and his valence level increased sharply at this time point. During the last time interval, valence of Subject 5 remained at a high level, meanwhile, stress of Subject 6 increased to the highest level of the whole conversion since he was thinking about the ways of implementing the proposed solution.

Thus, EEG-based emotion and stress recognition comparing to traditional questionnaire allows us to do data analyses during the task performance.



(a)



(b)

Figure 5. The changes of emotional state (a) and stress (b) averaged by 10 seconds over four 5-minute intervals using EEG-based recognition.

V. CONCLUSION

We designed and implemented an experiment to study the user experience when Skype conversation is augmented with haptic interaction. The user study was done in two ways: with traditional questionnaire and with novel EEG-based tools.

The EEG data were recorded during the experiment, and valence levels of emotions and stress levels were recognized and analyzed. The results of data analyses confirmed that active haptic interaction gives more satisfaction to the user over passive haptic interaction over Skype. We are planning to extend the experiment with more subjects to fully validate the hypotheses.

Although the applied subject-dependent emotion and stress recognition algorithms still need further improvement in accuracy, including implementing better artifacts removal algorithms, the proposed neurocognitive tools for human-computer interaction study open new possibilities in neuroscience based design. The novel tools could be applied for both hardware and software assessment in user studies and in human factor studies in different applications.

ACKNOWLEDGMENT

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centers in Singapore Funding Initiative, it is also supported by MOE Singapore Funding RG17/15.

REFERENCES

[1] S. U. Rehman, J. Sun, L. Liu, and H. Li, "Turn Your Mobile Into the Ball: Rendering Live Football Game Using Vibration," *IEEE Transactions on Multimedia*, vol. 10, pp. 1022-1033, 2008.

[2] Y. Kim, J. Cha, J. Ryu, and I. Oakley, "A tactile glove design and authoring system for immersive multimedia," *IEEE MultiMedia*, vol. 17, pp. 34-45, 2010.

[3] F. Danieau, J. Fleureau, A. Cabec, P. Kerbiriou, P. Guillotel, N. Mollet, *et al.*, "Framework for enhancing video viewing experience with haptic effects of motion," in *2012 IEEE Haptics Symposium (HAPTICS)*, 2012, pp. 541-546.

[4] M. Meehan, B. Insko, M. Whitton, and J. Frederick P. Brooks, "Physiological measures of presence in stressful virtual environments," *ACM Trans. Graph.*, vol. 21, pp. 645-652, 2002.

[5] Y. P. Lin, C. H. Wang, T. L. Wu, S. K. Jeng, and J. H. Chen, "EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, Taipei, 2009, pp. 489-492.

[6] K. Schaaff, "EEG-based Emotion Recognition," Diplomarbeit am Institut fur Algorithmen und Kognitive Systeme, Universitat Karlsruhe (TH), 2008.

[7] Z. Lan, O. Sourina, W. Lipo, and L. Yisi, "Stability of Features in Real-Time EEG-based Emotion Recognition Algorithm," in *Cyberworlds (CW), 2014 International Conference on*, 2014, pp. 137-144.

[8] Y. Liu and O. Sourina, "EEG-enabled Affective Human-Computer Interfaces," *UAHCI/HCII 2014, Part I, LNCS 8513 proceedings*, pp. 536-547, 2014.

[9] Z. Lan, O. Sourina, L. Wang, and Y. Liu, "Real-time EEG-based emotion monitoring using stable features," *The Visual Computer*, pp. 1-12, 2015.

[10] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, pp. 277-283, 6// 1988.

[11] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1175-1191, 2001.

[12] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, pp. 186-197, 2010.

[13] S. A. Hosseini, M. A. Khalilzadeh, M. B. Naghibi-Sistani, and V. Niazmand, "Higher order spectra analysis of EEG signals in emotional stress states," in *Information Technology and Computer Science (ITCS), 2010 Second International Conference on*, 2010, pp. 60-63.

[14] R. Khosrowabadi, C. Quek, K. K. Ang, S. W. Tung, and M. Heijnen, "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress," in *Neural Networks (IJCNN), The 2011 International Joint Conference on*, 2011, pp. 757-762.

[15] X. Hou, Y. Liu, O. Sourina, T. Y. R. Eileen, L. Wang, and W. Mueller-Wittig, "EEG based Stress Monitoring," in *IEEE International Conference on in Systems, Man and Cybernetics (SMC)*, 2015, pp. 3110-3115.

[16] Bradley M. M. and L. P.J., "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual," University of Florida, Gainesville2007.

[17] X. Hou, F. Trapsilawati, Y. Liu, O. Sourina, C. H. Chen, W. Mueller-Wittig, *et al.*, "EEG-based human factors evaluation of conflict resolution aid and tactile user interface in future Air Traffic Control systems," in *In 4th International Conference on Human Factors in Transportation*, accepted, July 2016.

[18] K. Stytsenko, E. Jablonskis, and C. Prahm, "Evaluation of consumer EEG device Emotiv EPOC," in *MEi: CogSci Conference 2011, Ljubljana*, 2011.