PRELIMINARY FINDINGS ON IMAGE PREFERENCE CHARACTERIZATION BASED ON NEUROPHYSIOLOGICAL SIGNAL ANALYSIS: TOWARDS OBJECTIVE QOE MODELING

Tiago H. Falk[1,4], *Yael Pomerantz*[2], *Khalil Laghari*[1], *Sebastian Möller*[3], *Tom Chau*[4]

[1] Institut National de la Recherche Scientifique, INRS-EMT, Montreal, Canada
[2] Universidad Iberoamericana, Mexico City, Mexico
[3]Telekom Innovation Labs, TU Berlin, Berlin, Germany
[4]Bloorview Research Institute, University of Toronto, Toronto, Canada

ABSTRACT

Image preference is a subjective factor which plays an important role in Quality-of-Experience (QoE) modelling. Traditionally, preference characterization has been quantified via questionnaires or subjective evaluations. Current advances in neurophysiological signal acquisition, however, have allowed for such "non-measurable" subjective parameters to be quantified objectively. In this pilot study, we explore the use of neurophysiological signals as correlates of image preference characterization. Experiments with seven participants have shown promising results and mental states associated with preferred and non-preferred images, as well as baseline neutral state could be classified with above-chance levels.

***Index Terms*—** QoE, NIRS, PANS, image preference.

1. INTRODUCTION

As argued by Winkler & Mohandas, multimedia Quality-of-Experience (QoE) encompasses various influencing (objective) parameters (e.g., measurable technical attributes) and subjective quality features (henceforth referred to as "quality features") [1]. Examples of influencing parameters can include display sizes and lighting conditions. Quality features, in turn, can include user preference, expectation, and emotional state [2]. Commonly, quality features are obtained via questionnaires or subjective evaluations which ask participants to rate their e.g., preference or acceptance levels in response to presented stimuli [3]. Perkis et al further argue that such subjective insights are "non-measurable" [4].

With recent advances in neurophysiological biosignal acquisition and analysis, there is growing evidence that indicators of quality features may indeed be quantitatively measured. As an example, cerebral blood flow, measured via near-infrared spectroscopy (NIRS), has been linked to different emotional states with positive/negative emotions being discriminated with up to 72% accuracy [5]. Additionally, physiological signals harnessed from the peripheral autonomic system (PANS) have also been explored for affective state characterization [6]. Recently, cerebral blood flow information was used to characterize subjective preference to different beverages [7]. In the pilot experiment described herein, we explore the usefulness of combining NIRS and PANS signals for automated image preference characterization. Image preference plays a key role in objective QoE modelling [8] and has been shown to also improve video quality measurement [1]. An experiment with seven participants showed that "preferred," "not preferred," and neutral baseline states could be quantitatively characterized with accuracy significantly better than chance. With advances in wearable computing, biosignal analysis should play a key role in objective QoE monitoring.

2. METHODOLOGY

Seven healthy subjects (aged 20.1 + 1.1 years, mean + SD) participated in this study. All subjects were recruited amongst students and staff at the Bloorview Research Institute, University of Toronto. Subjects were pre-screened for neurological, metabolic, cardiopulmonary, psychological, or drug- or alcohol-related conditions that could affect either the measurements or their ability to follow the experimental protocol. All subjects had normal or corrected-to-normal vision and provided signed consent for their participation in the study.

NIRS signals were collected using an Imagent Functional Brain Imaging System (ISS Inc, Champaign, IL, USA) with 7 light emitter pairs (one at a 690-nm wavelength and the other at 830-nm) and three photomultiplier tube detectors. A custom-made headgear made of polyurethane material was developed to accommodate the NIRS sensors (running at a sampling rate of 31.25Hz) against the forehead of the participants. To allow for consistent placement, the headgear was positioned on the participant's forehead above the eyebrows, with two detectors lying approximately over the FP1 and FP2 positions of the International 10-20 System, while the third detector lied above the nose line at the centre of the forehead. A complete description of the NIRS technology, and its use in cerebral blood flow and oxygenation monitoring, is beyond the scope of this paper and the interested reader is referred to [7] and the references therein. PANS signals, in turn, were collected with a ProComp Infiniti System (Thought Technology, Montreal, Canada) with the following sensors: blood volume pulse (BVP), galvanic skin response (GSR), respiration belt, and skin temperature; a sampling rate of 256Hz was used. The two systems were synchronously connected via hardware and triggered via an external trigger box.

The experimental protocol consisted of two parts. First, participants were asked to rate their preference for images falling under 10 categories: beverages, food, desserts, clothes, cars, accessories, movies, TV shows, purses, and shoes. This was done to avoid arbitrary decisions and to guarantee that preferred and non-preferred images were presented in a balanced manner during the second part of the experiment. During the second part, each participant was presented with 70 blocks of image pairs; each block lasted 45 seconds. Each block consisted of 10 seconds of baseline where participants were asked to relax, followed by a five-second window where participants were shown a copy of the two images to be presented, followed by two ten-second intervals where a preferred and non-preferred image was presented in random order. A block finished with a 10-second baseline interval and a scoring period for participants to report their preferences.

NIRS signals were first pre-processed using a wavelet filter to remove signal artifacts such as the so-called Mayer wave and further transformed to oxygenated hemoglobin and deoxygenated hemoglobin concentrations using the modified Beer-Lambert law (for more details, refer to [7]). Relative to the PANS signals, for the BVP sensor, the following features were extracted: number of peaks; mean, median, and standard deviation of upward and downward slopes. For GSR: signal maxima, minima, range, mean, slope, and peak value. For the skin temperature sensor: signal average and standard deviation. For the respiration sensor: number of breaths, signal average, standard deviation and respiration length line.

3. EXPERIMENTAL RESULTS AND DISCUSSION

In order to classify the neurophysiological data into the three classes (baseline, preferred image, non-preferred image), three conventional classifiers were explored, namely support vector classifier (SVC), neural network (NN), and a classification tree. Five-fold cross validation was used to partition the data into training and test sets and feature pruning was performed automatically by the classifiers. It was observed that the PANS signals conveyed the most useful information, with the NIRS-plus-PANS signal combination resulting in the best performance. Accuracy as high as 72% was obtained with a NN classifier and 71% with a classification tree; results were somewhat lower with SVM (67%). In all three cases, the obtained results were significantly different than chance at a 99.9% level. A high inter-subject variability was observed, with one participant resulting in 48% accuracy, which nonetheless, is still significantly different than chance at a 99.9% level (binomial t-test). Ongoing studies are focusing on developing participant-specific classifiers to compensate for this neurophysiological signal variability. A video preference task may also be used to elicit a stronger neurophysiological response, thus improving classifier performance in addition to providing time-varying insights. Ultimately, it is hoped that such an objective multimedia preference classifier can provide useful subjective aspects to a QoE model.

4. CONCLUSIONS

This pilot study has explored the usefulness of harnessing neurophysiological signals as correlates of image preference characterization. Conventional features were extracted from neurophysiological signals and used to train three different classifiers. An experiment with seven participants showed high inter-subject variability. Notwithstanding, classifiers obtained accuracies significantly greater than chance for all participants. These promising preliminary results suggest that "non-measurable" quality features may be measured via neurophysiological signals. Ultimately, it is hoped that these features will be incorporated into objective QoE models.

5. REFERENCES

[1] S. Winkler and P. Mohandas, "The evolution of video quality measurement: from PSNR to hybrid metrics," *IEEE Trans Broadcasting*, vol. 54, pp. 660–668, 2008.

[2] F. Pereira, "Sensations, perceptions and emotions towards quality of experience evaluation for consumer electronics video adaptations," in *Intl Workshop Video Proc Quality Metrics for Consumer Electr*, 2005.

[3] S. Jumisko-Pyykkö and T. Utriainen, "User-centered quality of experience: is mobile 3d video good enough in the actual context of use?" in *Intl Workshop Video Proc Quality Metrics for Consumer Electr*.

[4] A. Perkis, S. Munkeby, and O. Hillestad, "A model for measuring quality of experience," in *Proc. Signal Processing Symposium*. IEEE, 2006, pp. 198–201.

[5] S. Moghimi, A. Kushki, S. Power, A. Guerguerian, and T. Chau, "Automatic detection of a prefrontal cortical response to emotionally rated music using multi-channel near-infrared spectroscopy," *Journal of Neural Engineering*, vol. 9, p. 026022, 2012.

[6] R. Picard and J. Klein, "Computers that recognise and respond to user emotion: Theoretical and practical implications," *Interacting with Computers*, vol. 14, p. 26, 2002.

[7] W. Luu and T. Chau, "Decoding subjective preference from single-trial near-infrared spectroscopy signals," *Neural Eng*, vol. 6, p. 8, 2009.

[8] F. Kobayashi, M. Masuda, and T. Hayashi, "Qoe assessment method for mobile video services based on user motivation," in *Proceedings of SPIE*, vol. 8293, 2012.