Towards an implicit time-continuous assessment of Quality of Experience for immersive multimedia

Anne-Flore Perrin*, Touradj Ebrahimi*
*Multimedia Signal Processing Group, EPFL, Lausanne, Switzerland, {anne-flore.perrin, touradj.ebrahimi}@epfl.ch

***Abstract*—Quality of Experience (QoE) is a time-varying perceptual process. Despite its complexity and numerous impacting factors, recent efforts have been made to come up with subjective assessment methodologies of QoE. However, no time-continuous variant of subjective evaluation methodologies have been proposed. Considering that an explicit time-continuous assessment could impact the experience itself, validity of conclusions from such an evaluation can be put in question. The solution proposed here is an implicit time-continuous assessment methodology indicating the time-periods of an audio-visual stimulus during which an experience is assessed. The motivation behind this methodology is to provide additional and non-redundant information enabling further investigations in Sense of Presence (SoP). SoP is one of the major metrics in immersive technologies, highly correlated with QoE. This paper describes a work in progress for such an evaluation methodology design.**

I. Introduction

The design of subjective evaluation methodologies is complex as a multitude of influencing factors can prevent the validity of their conclusions. Standardization committees, such as the International Telecommunication Union (ITU), provide recommendations for subjective test methodologies.Explicit time-continuous scoring methodologies are described in the above-mentioned recommendations, such as asking subjects to vote continuously during display of a stimulus, using a slider. However, no implicit time-continuous assessment is proposed in those documents. On the other hand, recently, standardization committees have created new study groups in order to focus on implicit assessment methodologies using psychophysiological signals, such as brain activity (e.g. ElectroEncephaloGraphy (EEG)) as well as peripheral signals, such as heart activity (e.g. ElectroCardioGraphy (ECG)) and Electrodermal Activity (EDA). In recent studies, the investigation of the Sense of Presence (SoP), towards the assessment of Quality of Experience (QoE), was based on the analysis of multimodal modalities, such as self-reported scores and recordings of physiological signals. The former modalities define the ground-truth while the latter help in classification of the levels of presence experienced by subjects. However, the unrelated-activity noise contained in the physiological signals can bias the classification outcomes as it leads to inaccurate presence-related features. Reducing the unrelated-activity noise would help improving the reliability of the classification process. The envisioned method for noise reduction is to perform the feature extraction exclusively from phases throughout which subjects experienced presence. The identification of such phases lies in a time-continuous evaluation of presence. The relevance of having a non-explicit assessment process is explained by the fact that a self-rating will prevent the subject to fully experience the visual content. A time-continuous explicit assessment methodology will create a bias in the gathered information since the engagement of the subject will be focused on the evaluation instead of being dedicated to the stimulus consumption. This will infer a reduction of the level of presence felt, as this latter is a measure impacted by the combination of the level of immersion and involvement [1]. Presence is one of the main criteria of immersive technologies evaluations and is highly correlated with QoE. Henceforth, references to the SoP imply a potential extension of the results to the QoE.

The developed implicit time-continuous assessment methodology is premised on the assumption that, when experiencing presence, one's self conscience, of body and behavior, in physical environment is decreased leading to a different kind of self awareness state. In many situation this is combined with a state of relaxation. The measure of the relaxation state of a subject is challenging and can result from various processes. The proposed process attempting such a measure consists in requesting the subject to constinuously tighten a physical ball at a same level of pressure while recording subject's hand ElectroMyography (EMG). It is expected that the variations of subjects' relaxation state are indicated by EMG signals, leading to the identification of phases of presence

The scope of this paper is limited to controlled environment evaluations. Further research and other solutions will be needed to be envisioned for uncontrolled environments. In addition, this method is developed for immersive multimedia, and cannot be applicable in all contexts.

The remainder of the paper is organized as follows. Related work is described in section II. The objectives of depicted idea are detailed in section III. A precise definition and justification of the time-continuous assessment methodology is provided in section IV. It also contains the constraints and the improvements as well as the future questions to be solved in order to further develop this idea. The current work under progress is presented in section V. Finally, conclusions are drawn in section VI.

II. Related Work

Subjective test methodologies consist in asking the appraisal of the subjective perception of thoughtfully selected stimuli to a sample of a population. Self-assessments of test participants are conscious responses, impacted by perceptual and cognitive processes. In order to have more insights into the construction of the subjective appraisal process, it is necessary to have spontaneous and less controlled responses to stimuli, as provided by physiological signals. Physiological signals can be divided into two categories: those originating from the peripheral nervous system (e.g. heart rate, EMG, Galvanic Skin Response (GSR)), and those coming from the central nervous system (e.g. EEG). Since a decade, attempts to include recordings of such signals in subjective tests procedures have been carried out [2]. When combining EEG with peripheral signals, the complementariness and redundancy of the information across modalities provide an efficient and robust measure of human responses [3]. Standardization committees are taking an interest in such evaluations as the ITU and Video Quality Experts Group (VQEG) have established working groups in order to develop recommendations for physiological assessment methodologies, such as P.PHYSIO, a 2018 ITU-T SG12 work program for speech processing[1].

Such committees have already provided time-continuous assessment recommendation consisting in using a continuous desk-mounted slider [4], [5]. Various time-continuous methods were also investigated in [6] for (semi) living-labs environments, comparing the slider method with new time-continuous assessments such as finger distance, finger count and the acceleration and orientation of a mobile phone. A specific attention has been devoted to user's opinion, the distraction created by an explicit assessment, the precision of the rating process and the reaction and scoring time. Most criteria in evaluation methodologies mentioned above are also relevant for the development of an implicit assessment. In the context of assessment of SoP, one does not experience presence all of a sudden but through a progressive feeling of immerssion and engagement, which result in a gradual increase of the level of presence during a certain period of time. Thus, SoP and QoE are lengthy processes, which should be accounted for in measuring reactions.

To the best of authors knowledge, there is no implicit time-continuous multimedia quality assessment, other than a thorough analysis of the entire physiological signals. The idea proposed here aims at an implicit time-continuous assessment of QoE by identifying phases of presence during consumption of multimedia content. The detection of phases of presence enables the reduction of presence-unrelated features in the signals to be processed and at the same time will provide more insights on the actual formation and experience of presence by indicating the duration and number of phases of presence.

III. Objectives

Experiencing presence is not a binary process pointing out whether or not a subject felt immersion while consuming a multimedia content. This explains the need of recording physiological and peripheral signals in order to be able to evaluate to which extent presence was experienced [7].

[1]http://www.itu.int/itu-t/workprog/wp_item.aspx?isn=13800

In the work conducted in [8], the processing of physiological and peripheral signals consisted of pre-processing signals (e.g. blinking artifacts and noise reduction), followed by feature extraction. A fusion operation was then applied on the features extracted from the various modalities. A vector formed by fusion of features, combined with the ground truth provided by self-assessments, was used to classify the results into different levels of presence.

The main objective of developing an implicit time-continuous assessment of SoP is the identification of *phases of presence*. A phase of presence is a time period during which the subject is immersed and demonstrates an involvement (moderate to high level of attention). The number of phases of presence, combined with their duration, indicate during which portion of a stimulus the subject experienced presence.

Those new features integrated in the above-mentioned classification process, will enable a more accurate feature extraction, solely based on presence-related signals. The expected results of the described process is the improvement of the measurement system in terms of robustness and reliability leading to a thorough understanding of the SoP.

It should be mentioned that in the context of this work, the analysis of physiological and peripheral signals provides a qualitative assessment of the level of presence. The detection of phases of presence is a quantitative assessment, providing new information, missing from previous studies.

In addition, an advantage of a time continuous assessment of QoE is the possibility of running tests on stimuli with long-duration. Self-assessment is usually required after each stimulus, providing an overall rate per stimulus. This process is validated for short-duration stimuli. Indeed, the longer a stimulus, the higher is the cognitive load on the subject. This leads to a measure of perception less reliable due to a higher number of cognitive processes impacting the assessment.

IV. Developed System

The idea developed here is based on the assumption that, during an immersive multimedia experience, one's physical self awareness decreases when presence increases. It has been observed in previous studies that the higher the level of presence experienced, the lower is the conscious of one's surrounding physical environment [8]. The assumption is the extension of this observation to the reduction of one's self awareness of physical behavior while experiencing presence.

The following describes how we propose to measure the conceptual constructs above-mentioned. If a subject is required to grasp firmly a ball at a constant pressure level, it is expected that hand pressure will slowly and steadily decrease after a phase of presence has been reached. It should be emphasized that subjects should not be aware of the purpose of the EMG signals recordings. The justification of such a decision is that the assessment should be implicit. A subject aware of the aim of the pressure signals will probably consciously or unconsciously interfere and bias the recordings.

The start of a phase of presence is detected when the relative pressure decreases to hit a predefined threshold. A sudden increase in one's hand pressure will result from the subject's change of state of consciousness when he/she remembers the instruction of grasping the ball, indicating the end of a phase of presence. The detection of the beginning of a presence phase will be far more challenging than the identification of its end.

The exact intensity of the pressure applied on the ball is not as the important as its variation. The EMG is a neuromuscular monitoring technique, recording the electrical activity produced by skeletal muscles. The analysis of such peripheral signals provides insights into the activation level of the muscles. The variation of the activation level of hand muscles is an indicator of the hand pressure.

Thus, the proposed system, aiming at identification of phases of presence, consists in recording EMG signals, measuring the activation level of muscles of the hand, when requiring subjects to tightening a ball at a constant pressure. The non-invasive EMG procedure records the potential difference between two electrodes. The observed signals are limited to superficial muscle activities recordings due to the placement of the electrodes on the surface of the hand.

There are several issues to tackle when designing such a system. For this work, we have to:

1) Select the electrode placements on the hand for an optimized measure of muscle activities. Only one signal should be sufficient to detect phases of presence.
2) Verify whether the variations of muscle activities are equivalent to the variations of pressure.
3) Define the processing to apply on the signals for identification of phases of presence. This includes the definition of the signal processing to apply on EMG recordings as well as of the detection process for phases of presence (e.g. investigation of the threshold indicating the start of a phase of presence).
4) Characterize the features extracted from the processed EMG signals and the fusion process to merge the new features with features of other physiological signals.
5) Evaluate the construct validity. The measurement should correspond to what is meant to be measured.
6) For each of the previous steps, design and run pilot tests in order to verify or appraise the efficiency of selected methods.
7) Determine or estimate if the assumptions hold. In this context, there are two main assumptions: (1) Relaxation states can be identified through the analysis of the hand muscles activities when one is asked to firmly grasp a ball at a constant pressure, (2) Those relaxation states can be assimilated to phases of presence.

During the verification of the construct validity of the test, several important points should be examined.

The modality of recording hand muscle activities could have been performed in several other manners. In addition, recording hand pressure is one among several other possibilities that could be envisioned for detection of phases of presence. The focus on this method is mainly due to the accessibility of the material. Once this method is shown promising, other methods could be developed and evaluated in order to select an optimal solution.

The signals could be too noisy or could not be sufficiently precise to show the muscle activities as the variation is supposed to be really slow and possibly barely noticeable. The observation of the variations in hand pressure (e.g. constant pressure, slight reduction, abrupt increase) through EMG signals has to be verified.

Variations of the hand pressure could be due to the tiredness of the subject when tightening the ball as well as muscles fatigue. Several pilot tests, especially regarding long stimuli assessment, will be required. Recommendations regarding the duration of an assessment and the overall duration of a test session would be necessary.

Requesting the subject to grasp firmly the ball will introduce a bias in EMG recordings. Some test participants may focus too much on the instruction. If so, it is more likely too few or no phases of presence will be detected from such participants' signals, making those subjects outliers of the study.

The extension of this subjective methodology assessing the SoP towards the evaluation of QoE will also require a lot of work, but it is out of the scope of this paper.

V. Work in progress

This ongoing work is in its infancy. This section describes the progress achieved to date, and the next step ahead.

The potential observation of variations in hand pressure measured from the EMG signals has been investigated. EMG signals from muscle activities translate pressure information to electric potentials in either positive or negative voltage. It is a common procedure to analyze only the positive values of such signals [9]. Moreover, the level of muscle activation is indicated by the amplitude of the signal.

Two electrodes Ambu blue sensor N-00-S[2], high-quality Ag/AgCl electric field sensors, were used combined to a Geodesic EEG System (GES) 300 [3] in order to record, amplify, and digitize the EMG signals. Signals were recorded, following the recommendations provided in [10], when investigating various types of pressure variations, such as constant level pressure and steadily decreasing pressure. Due to the inherent noise in electronic components, a high-pass filter was applied on the signals based on its power spectrum. It results that EMG signals enable analysis of variations in hand pressure. Figure 1 presents the EMG signals resulting from the application of different variation of pressure when tightening the ball.

Various electrodes placements were tested. The results presented above are from the placement of the electrodes on the extremities of the Abductor Pollicis Brevis (APB) muscle of the left hand of a right-handed subject. Future tests should consider the placement of the electrodes on the dominant hand of the subject. Indeed, handedness indicates the likeliness of

[2]http://www.ambu.co.uk/ukca/products/patient_monitoring_and_diagnostics/product/ambu_bluesensor_n-prod4183.aspx

[3]https://www.egi.com/clinical-division/clinical-division-clinical-products/ges-300

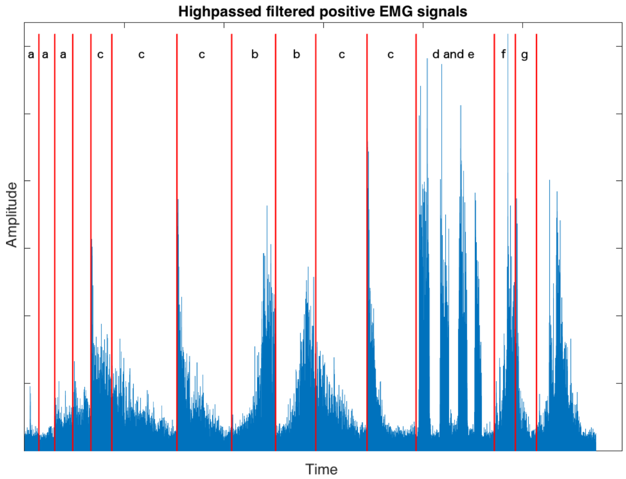

Fig. 1: High-pass filtered positive EMG signals and variants of applied pressure when grasping a ball.
with (a) constant level pressure, (b) steady increase in pressure, (c) steady decrease in pressure, (d) sudden increase in pressure, (e) sudden decrease in pressure, (f) step by step increase in pressure and (g) step by step decrease in pressure

better performance as a function of the hand used. A theoretical investigation followed by an experimental verification of the optimal position of EMG electrodes should be performed in order to reach more reliable and relevant signals. Based on an anatomical analysis of the hand, muscle position, depth of the subcutaneous tissues as well as the functioning of muscles electrical activities will be considered to recommend best electrodes placement positions. The selection of an electrode's position will be based on the acquired signals during a pilot test (e.g. signal-to-noise ratio, distortion of the signal and observation of hand pressure variations). In parallel of this process, an exploration of EMG signal processing enabling the identification of phases of presence will be carried out.

VI. Conclusion

This paper tackles the lack of implicit time-continuous subjective test methodologies providing the required information towards a better understanding of SoP. Previous studies took advantage of an overall implicit assessment of the level of presence experienced provided by physiological signals. However, no insights regarding the actual presence experienced at a time t of an immersive multimedia stimulus is provided by such studies. The described idea aims at the identification of phases of presence by using EMG signals. It is assumed that phases of presence can be assimilated to relaxation states of the subject resulting from reduction in physical behavior self awareness.

The wide range of different issues to tackle includes establishment of an assessment procedure (e.g. its design and validity evaluation) and the appraisal of the process accuracy and efficiency.

The results obtained so-far show that hand muscles activities enable identification of variations in hand pressure. This is promising as based on the designed assessment methodology, variations in hand pressure point out subjects' relaxation states, allegedly indicators of phases of presence.

The context used for this work is the improvement of SoP experiences classification using physiological signals. Nevertheless, numerous use-cases need such a time-continuous assessment methodology, which gives even more importance to this ongoing work.

Acknowledgment

The work behind this paper was possible thanks to European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 643072 - Network QoE-Net. Authors especially acknowledge funding received for these projects from Swiss State Secretariat for Education, Research and Innovation SERI.

References

[1] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence: Teleoperators and virtual environments*, vol. 7, no. 3, pp. 225–240, 1998.

[2] U. Engelke, D. P. Darcy, G. H. Mulliken, S. Bosse, M. G. Martini, S. Arndt, J.-N. Antons, K. Y. Chan, N. Ramzan, and K. Brunnstrom, "Psychophysiology-based qoe assessment: A survey," *IEEE Journal of Selected Topics in Signal Processing*, 2016.

[3] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, *Emotion Assessment: Arousal Evaluation Using EEG's and Peripheral Physiological Signals*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 530–537. [Online]. Available: http://dx.doi.org/10.1007/11848035_70

[4] *Methodology for the subjective assessment of the quality of television pictures.* ITU-R BT.500-13, January 2012.

[5] *Subjective video quality assessment methods for multimedia applications.* ITU-T P.910, April 2008.

[6] S. Buchinger, W. Robitza, M. Nezveda, E. Hotop, P. Hummelbrunner, M. C. Sack, and H. Hlavacs, "Evaluating feedback devices for time-continuous mobile multimedia quality assessment," *Signal Processing: Image Communication*, vol. 29, no. 9, pp. 921–934, 2014.

[7] A.-F. Perrin, M. Řeřábek, and T. Ebrahimi, "Towards prediction of sense of presence in immersive audiovisual communications," *Electronic Imaging*, vol. 2016, no. 16, pp. 1–8, 2016.

[8] A.-F. N. M. Perrin, H. Xu, E. Kroupi, M. Řeřábek, and T. Ebrahimi, "Multimodal dataset for assessment of quality of experience in immersive multimedia," in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, pp. 1007–1010.

[9] M. Reaz, M. Hussain, F. Mohd-Yasin *et al.*, "Techniques of emg signal analysis: detection, processing, classification and applications," *Biological procedures online*, vol. 8, no. 1, pp. 11–35, 2006.

[10] S. Day, "Important factors in surface emg measurement," *Bortec Biomedical Ltd publishers*, pp. 1–17, 2002.