

## A WEB BASED SUBJECTIVE EVALUATION PLATFORM

*Benjamin Rainer, Markus Walzl, and Christian Timmerer*

Multimedia Communication (MMC) Research Group, Institute of Information Technology (ITEC),  
Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria, *firstname.lastname@itec.aau.at*

### ABSTRACT

Preparing and conducting subjective quality assessments is a time consuming and expensive task. Therefore, we present a Web-based evaluation framework which aims on reducing the time needed for planning and designing a subjective quality assessment. The presented framework can be used for both crowdsourced and laboratory experiments. It should ease the task of designing a subjective quality assessment by providing a flexible framework. The framework has proven its applicability and flexibility to design and conduct assessments in the past and is available as open source.

**Index Terms**— Evaluation Platform, Crowdsourced Quality Evaluation, Laboratory Quality Evaluation, Quality Assessment Framework

### 1. INTRODUCTION

Planning, preparation, and conducting subjective quality assessments in the area of Quality of Experience (QoE) is time consuming and in many cases an expensive task. The ITU provides recommendations on the methodology and design of subjective quality assessments for different use cases, e.g., multimedia applications or television pictures, but mainly for experiments in the lab [1][2]. In order to reduce the costs of laboratory subjective quality assessments, a new trend using *crowdsourcing* has evolved recently. The idea behind crowdsourcing is to conduct subjective quality assessments online such that participants may only need an up-to-date Web browser and an Internet connection to participate in the experiment. In contrast to the well-known laboratory experiments where the experiment is conducted in a controlled environment, crowdsourced experiments are not controlled and, thus, may result in cheating, unintended environment, no possibility to ask questions, etc. [3]. The participants may be acquired – in exchange for a small compensation – from different crowdsourcing platforms such as Amazons' Mechanical Turk or Microworkers.

In [4], *QualityCrowd* is presented, a platform designed for Mechanical Turk. Another platform for conducting crowdsourced QoE assessments is presented in [5] which is tailored to use cases where pair comparison is the preferred choice of the evaluation method.

In this paper, we present our subjective evaluation platform which is designed for both laboratory and

crowdsourced experiments and allows for an easy design and configuration of a subjective quality assessment. Furthermore, it is designed in such a way that it can be used for any existing crowdsourcing platforms and is intended to reduce the amount of time needed for designing and conducting a subject quality assessment.

### 2. WEB BASED EVALUATION PLATFORM

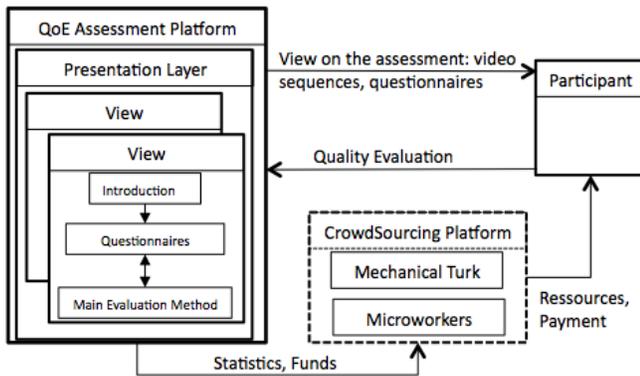
Our evaluation framework is based on the recommendation of the ITU for subjective quality evaluations of multimedia applications and television pictures [1][2].

The framework was mainly developed to conduct subjective quality assessments in the domain of sensory effects and has been successfully used to conduct subjective quality assessments in this domain. For example, in [6], our platform was used to investigate the influence of sensory effects on the QoE and the emotional response.

The proposed evaluation platform requires an HTTP server with PHP support and a MySQL database. It can be integrated into any crowdsourcing platform as long as there is the possibility to embed external Web sites into a crowdsourced task. For example, both Mechanical Turk and Microworkers support this possibility. The evaluation framework is open source and available at [7].

#### 2.1. Architecture

The evaluation platform itself follows the Model-View-Controller (MVC) pattern. Figure 1 illustrates an overview of the proposed evaluation framework which was developed in HTML and PHP. The introduction and questionnaires can be configured separately from the test methodology and allow for including control questions during the main evaluation. The voting possibility can be configured independently from the test methodology; thus, providing more flexibility in selecting the appropriate voting mechanism and rating scale. The predefined voting mechanisms include the common HTML interface elements and some custom controls like a slider in different variations. This offers more flexibility in the selection of the rating scale and the interface elements for voting possibilities in comparison to [4]. The platform consists of a management layer and a presentation layer. The management layer allows for maintaining the user study such as adding new questions



**Figure 1. Overview of the Evaluation Framework.**

or multimedia content and setting up the test method to be used (e.g., single stimulus, double stimulus, pair comparison, continuous quality evaluation). Therefore, this platform supports the most used test methods for conducting subjective quality assessments in comparison to [4] and [5], which support only a subset or in the case of [5], only a single test method. The presentation layer is responsible for presenting the content to the participants. This allows providing different views on the created study and, thus, one can define groups to which the participants may be randomly assigned (or in a defined way). After a participant has finished the user study, the gathered data is stored in the MySQL database. The platform provides means for an easy data analysis by developing PHP scripts, that read and analysis the data from the MySQL database.

Furthermore, the platform offers methods of tracking the participant's behavior during a subjective quality assessment by detecting if the Web browser's window has the focus and by calculating how much time each participant needed for each stimuli presentation and voting phase.

## 2.2. Stimuli Presentation

The stimuli presentation can be configured independently from the test method. The stimuli presentation may be combined with the voting possibility to support continuous quality evaluations. Furthermore, plug-ins may be embedded for laboratory experiments where external hardware is used. Multimedia content is presented by using either the HTML5 *video* element or the Adobe Flash Player. The decision which technique is used to render multimedia can be set through the management layer if one of them is needed explicitly. Otherwise, the decision is based on the codecs used and which of them does the client's Web browser support.

Furthermore, there is support for adding JavaScript to alter the presentation of the stimuli or for any other purpose.

## 3. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a subjective quality evaluation framework, which has been successfully used to conduct subjective quality assessments in the domain of sensory effects and which reduces the time needed for designing and conducting subjective quality assessments. Furthermore, the framework can be easily extended and can be adapted to meet the requirements for a subjective quality assessment in the field of multimedia. Additionally, with the use of HTML5 and the Adobe Flash Player it supports a broad range of codecs on various browsers.

Future work includes the improvement of the management layer (e.g., allowing an easier configuration) and additional monitoring techniques like the monitoring of the mouse motion.

**Acknowledgments:** This work was supported in part by the EC in the context of the ALICANTE (FP7-ICT-248652), SocialSensor (FP7-ICT-287975), and QUALINET (COST IC 1003) projects and partly performed in the Lakeside Labs research cluster at AAU.

## 4. REFERENCES

- [1] ITU-R Rec. BT.500-13, "Methodology for the subjective assessment of the quality of television pictures", 2012.
- [2] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications", 2008.
- [3] T. Hossfeld, M. Seufert, M. Hirth, T. Zinner, P. Tran-Gia, R. Schatz, "Quantification of YouTube QoE via Crowdsourcing," *2011 IEEE Int'l Symposium on Multimedia (ISM)*, pp. 494-499, Dec. 2011.
- [4] C. Keimel, J. Habigt, C. Horch, K. Diepold, "QualityCrowd - A Framework for Crowd-based Quality Evaluation," *Picture Coding Symposium (PCS)*, pp. 245-248, May 2012.
- [5] K.-T. Chen, C.-J. Chang, C.-C. Wu, Y.-C. Chang, C.-L. Lei, "Quadrant of Euphoria: A Crowdsourcing Platform for QoE Assessment", *Network*, IEEE, vol. 24, no. 2, pp. 28-35, Mar. 2010.
- [6] B. Rainer, M. Waltl, E. Cheng, M. Shujau, C. Timmerer, S. Davis, I. Burnett, C. Ritz, H. Hellwagner, "Investigating the Impact of Sensory Effects on the Quality of Experience and Emotional Response in Web Videos", *Proc. of the 4th Int'l. Workshop on Quality of Multimedia Experience (QoMEX'12)*, Yarra Valley, Australia, pp. 278-283, Jul. 2012.
- [7] Sensory Experience Lab, <http://selab.itec.aau.at>. Accessed 30.01.2013.