4th International Workshop on Perceptual Quality of Systems (PQS 2013)
2-4 September 2013, Vienna, Austria

Wizard of Oz Approach for a 3DVO Video Communication System

Sara Kepplinger[1], *Denise Tobian*[1]

[1]Institute for Media Technology, Technische Universität Ilmenau, Germany
Sara.kepplinger@tu-ilmenau.de, denise.tobian@tu-ilmenau.de

Abstract

This work describes the concept and the test of a Wizard of Oz Study approach. This is used in order to overcome the lack of interactivity of a pre-prototypical video communication application. The video communication system solves the problem of missing eye-contact by using free viewpoint video technology. Thereof, the quality should be evaluated. However, the system is not yet applicable in real-time due to extensive rendering. Therefore, a simulation in respect of quality evaluation is intended. The research question is towards which extent the test participants do not recognize the simulation and how this profits the design and evaluation phase. The results show the applicability of the Wizard of Oz Study approach in this context as well as critical issues on social interaction which needs to be taken into consideration and which we want to discuss within the workshop.

Index Terms: Wizard of Oz Prototyping, QoE, video communication, eye contact, 3DVO

1. Introduction

Free viewpoint video objects (3DVO) technology is used in order to overcome the lack of natural eye contact in video communication systems (see: [1]). However, this concept is not yet real-time capable because of extensive video processing. In order to integrate evaluation activities and their results (see: [2]), interactivity is seen as an important factor in this case [3]. The research aim is to investigate how the evaluation activities can be continued and include the factor interactivity although the video communication system does not support it at the moment.

This paper presents an approach in order to allow subjective quality assessment already without real-time capability including the factor interactivity and the real usage situation. Therefore, the Wizard of Oz method was used exemplarily in order to define whether this might be a useful method. The question is towards which extent the test participants do not recognize the simulation and how this profits the design and evaluation phase. As background information, the approach of eye contact support via 3DVO and subjective assessment of quality in this context are presented in a first step.

2. Background

There are several approaches available working on eye contact support in video communication and defining technical limits (as summarized in [4]). Based on the approach by Weigel and Treutner [1] this work tries to overcome the lack of real-time functionality of such approaches in order to facilitate subjective assessment of quality in this context and allow further developments of the system based thereon. There are other video communication system approaches available which are real-time capable (e.g. [5]). However, by Weigel and Treutner [1] presented concept of eye contact support is providing a different view point within the same usage context.

Several approaches on quality evaluation consider the influence of usage contexts (e.g. [6]). This usage context is not only becoming important when it comes to overall quality definition. The present system which is not yet real-time capable has a defined usage context but may vary (e.g., video communication on a mobile, on the run, at a stationary system, at home). The authors intend subjective quality assessment of the video communication system presented by Weigel and Treutner (see: [1]). A concept on subjective quality assessment therefore was presented in 2011 (see: [7]). Herein, there are different challenges described apart from missing real-time functionality. Namely, visual quality-dependent issues like occlusions or holes. In order to rate the influence of these challenges on the perceived quality, adequate evaluation methods taking into account the real usage situation and the system's inherent purpose of *interactivity* are necessary.

3. Method

There are different approaches available in order to use prototypes. For example low-fidelity prototypes like storyboards or paper prototypes which are easy to realize, but do not consider details, or moving pictures. Another approach is high-fidelity prototyping which is already very similar to the developed system but does not realize all functions. This kind of prototyping is envisaged with the Wizard of Oz method.

The Wizard of Oz method allows the interaction between human and system without expensive computer support. Herein, essential activities are done by a person (i.e. the operator called Wizard) [8]. This technique is successfully applied in situations where the system itself is not ready. The investigation of interactivity and its accompanying influencing factors is important within the design process. For example, in order to test the user experience with robots [9]. However, there are challenges and principles (e.g., realistic rule based reactions, hiding the wizard) which have to be considered when designing a Wizard of Oz study (see also [10]). This is especially true within the usage context *dialog*, wherein the challenge of naturalness (i.e. allowing the impression of a realistic conversation) plays a big role [11], [12]. Pitfalls like unplanned variations within conversations have to be investigated.

This work presents a concept on the usage of Wizard of Oz including a test, its results, and suggestions for an improved concept.

10.21437/PQS.2013-28

3.1. Applied Wizard of Oz Prototyping Concept

Two rooms were used for the study, one for the participants and one for the wizard. A display (46 inch) and a notebook as player devices were used for the presentation (see Figure 1). Thereon, a Skype [13] video call application as well as the simulation application was used.

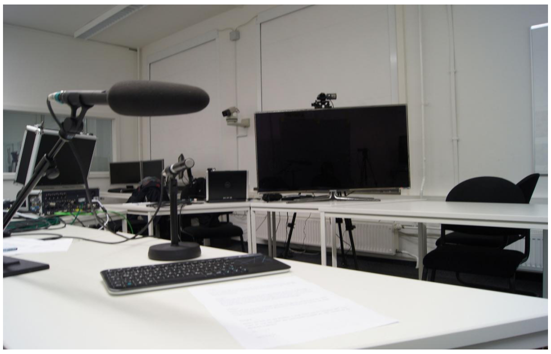

Figure 1: *Test setup from the test participants' position*

The simulations were recorded with a resolution of 640 x 360 pixels with 25 frames per second in order to guarantee a real-time like playback. These dialogs were recorded in the same room with same conditions (e.g. setting, clothes, light) where the real-time dialogs were done by the wizard during the test.

These are one minute long and contained dialog topics like a private removal (dialog 1), world travel (dialog 2), business dialog about a website design (dialog 3), and the talk to a technical support because of a broken computer (dialog 4). The length of the dialogs was defined due to the consideration of time until the simulation cannot be discovered and based on suggestions by standards [14] (whereas herein the suggestions are rather related to transmission influenced audio quality and therefore ask for at least two minutes length). The prerecorded test sequences can be requested via E-Mail to sara.kepplinger@tu-ilmenau.de. Figure 2 presents the wizard which is at the same time the conversation partner.

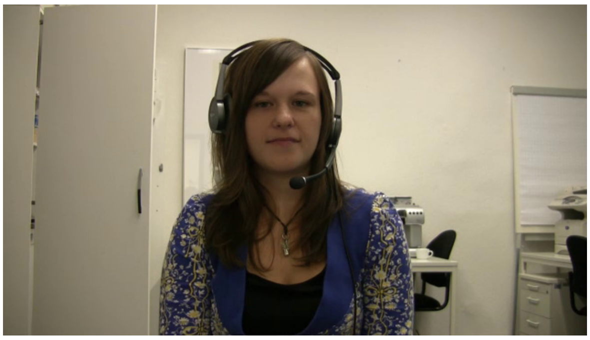

Figure 2: Snapshot of a prerecorded dialog

The concept of the dialogs allowed the test participants to act as a conversation partner to give some answers or input.

Reacting to this input, the wizard simulates answers with pre-recorded and within the real-time conversation with pre-defined reactions. Figure 3 shows the setup from the wizard's point of view.

For the simulation an application is created with Adobe Flash and LabVIEW [15] after a detailed storyboard and operating plan.

Figure 3: *Wizard's test room*

Herein, the application starts with a picture of a desktop which looks exactly the same as the desktop used within the real video communication. Then the test participant is able to answer an incoming call via click on the green button. Due to this action the intended video sequence is presented. The video sequence is followed by a looped other sequence during the time the test participant is allowed to answer. After that, the wizard initializes the next fitting sequence to present. The next sequence to choose is pre-defined within a text based document accessed by the Flash program. These text based documents may be changed by the wizard using LabVIEW, if necessary (see Figure 4).

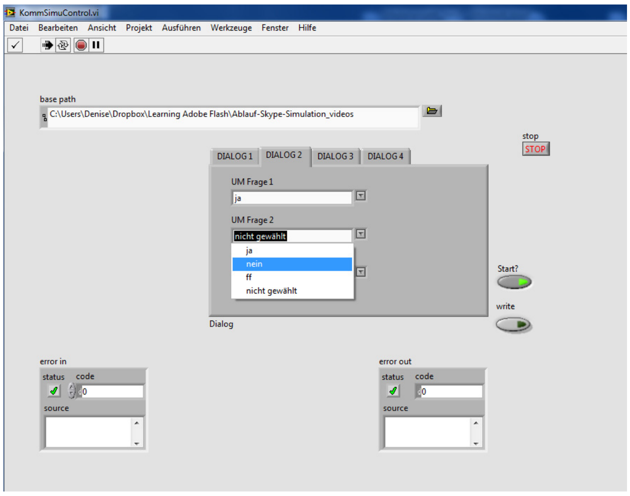

Figure 4: *Wizard's user interface to play videos*

First of all, four dialogs were presented using the simulation. In a next step, the application was changed and the next two dialogs (no. 5, 6) in real-time with Skype were presented. This change was concealed with a pretended short lasting "technical problem". After every dialog the participants were asked for their ratings regarding the quality of the dialog.

At the end of the test, the participants were interviewed to check if they recognized the simulation and if they saw some differences in the dialogs. With this, it was possible to get an idea if and where the simulation needs to be optimized.

The test participants were recruited based on pre-experience with video communication. Another prerequisite was that the participants should not be experienced in quality assessment of visual representations, as we did not want to pay attention on qualitative differences.

4. Results

To explore if the simulation under usage of the Wizard of Oz prototype is possible, 7 participants (2 men, 5 women) took part at the test. The participants were between 24 and 27 years old, the average age of the participants was 25.5 years. All of them had previous experience with video communication, mostly in a private context, whereof 5 out of 7 use video communication three times per month or less often.

The interviews showed that all dialogs seemed to be unnatural to the participants, once because it was no real face-to-face conversation, and once because of the predefined script. However, the majority of the participants pointed out that they realized the simulation only within the interview (majority) or as soon as they got the direct comparison by switching to the pre-recordings to the pre-defined live conversation or vice versa (two participants). One participant suspected a simulation from the beginning on.

Furthermore, the subjects did not realize differences regarding to the eye contact even though half of the dialogs had eye contact and the other half had no eye contact. Once there was no eye contact provided, the participants observed other parts in the picture, mostly the mouth of the conversation partner. Five out of seven participants indicated that the audio quality is more important than the video quality.

Figure 5 shows the quality rating of the above described different dialogs on a 0-100 scale per participant.

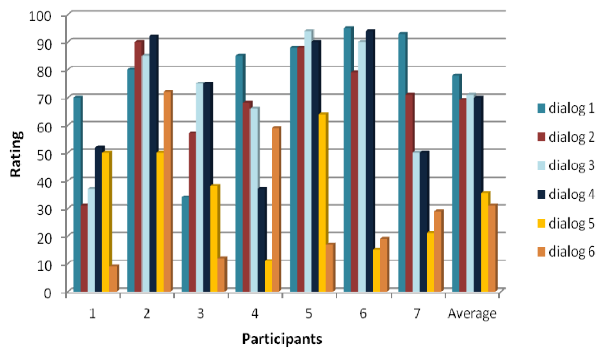

Figure 5: *Ratings of dialog 1-6*

As also seen in Figure 6, dialog 1 got the best mean score for overall quality (77.9%) by all test participants. Herein, five participants rated the overall quality of this dialog between 80% and 93% (average score). Dialog 2 to 4 got mean scores between 69% and 71%, but the participants rated these dialogs more differentiated. The two dialogs (no. 5 and 6) via the real video communication got a worse rating (although both, the simulated and the Skype conversations followed the same content wise script). Dialog 5 got a mean score of 35% (with ratings between 11% and 64%) and dialog 6 only 31% (with ratings between 9% and 72%).

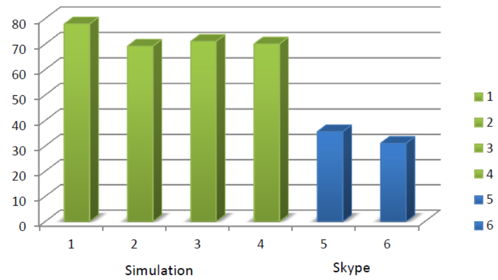

Figure 6: *Mean scores of overall quality of dialogs 1-6*

The distribution of the differentiated ratings for dialog 5 and 6 can be described by the differentiated quality of the used video communication system. The low quality of the real-time conversations (no. 5, 6) might depend on the content-wise repetition as well as on the transmission effect given by the internet connection (which was not necessary for the simulation). The delays based on updating the files during the simulation to start a new video sequence were not perceived as disturbing. A significant influence based on the conversation topics could be excluded.

Nevertheless, conclusions out of the comparison of Skype and simulated conversation situations can be made, as the quality of the simulations was rated as rather good with 70% average. This leads to the conclusion that possible discrepancies recognized by the test participants within the dialogs may not have a severe influence on the rating of quality when only using the simulation. This means, that participants would accept these simulated dialogs within a video communication.

Although it is not possible to compare both conditions, we got insight on the way how simulated dialogs are perceived and assessed. Paying attention to only simulated dialogs, Figure 5 shows that three out of seven participants rated the quality of the four simulated dialogs rather similar (Participant 2, 5, and 6). Herein, the range of the single ratings is between 6% and 16%. The ratings of the other participants (1, 3, 4, and 7) are more spread, and within a range between 39% and 43% of one test person. This shows that even recordings of dialogs, which are pre-defined as rather similar visual quality, are rated differently by different participants. This might change if a more variable range of quality is presented for quality assessment as usually done within standardized subjective quality evaluations.

In conclusion, 3 out of 7 participants suspected that some dialogs were realized with a simulation and 4 of the participants were surprised about this fact. The most influencing factor leading to the recognition of the simulation is the transition or editing between the single prerecorded video sequences. Generally speaking, the used setup of the prototype seems to be feasible in order to simulate video conversations. Herein, in a follow up study a quality assessment during the video conversation simulation using prerecorded sequences visualizing 3DVO can be considered. However, already the small amount of data collected indicates, that careful preparation of the dialogs for the simulation, as well as the effect given by the transmission for real-time applications, have to be considered when conducting a subjective quality assessment study on overall quality and eye contact perception.

5. Summary and Outlook

A concept and the test of a Wizard of Oz video communication prototype are presented. This approach seeks to overcome the lack of interactivity of a video communication application using 3DVO for eye contact support in order to conduct subjective quality assessments with a not real-time capable system but considering the influencing factor *interactivity*. The challenge of simulating *dialog* with a pre-defined script has to be investigated further, as it supports a feeling of unnaturalness. However, free speech within this approach is not possible (due to the real-time inability).

In order to investigate the impact of transmission further and to pay attention to the differentiated ratings between the simulation and the live Skype conversation (based on a pre-defined story), it is suggested to do this experiment with either more stable transmission conditions or to transmit the simulations in the same way as the other dialogs in the future.

The question was towards which extent the test participants do not recognize the simulation and how this profits the design and evaluation phase. Our implication towards this question based on presented study is, that if there is no switch between simulated and real-conversation a simulation may be used for the quality assessment purpose to investigate interactivity. However, this asks for the investigation of critical issues like very natural dialog concepts.

6. References

[1] Weigel, C. and Treutner, N., "Establishing eye contact for home video communication using stereo analysis and free viewpoint synthesis", Proc. IS&T/SPIE Conf. on Electronic Imaging 2012, San Francisco, CA, USA, (EI 2012), 8290-2.

[2] Kepplinger, S., "Linking quality assessment of free-viewpoint video objects up with algorithm development", Proc. SPIE-IS&T Conf. on Electronic Imaging 2012, San Francisco, CA, USA, (EI 2012), 8293-29.

[3] Rittermann, M., "Zur Qualitätsbeurteilung von 3D Videoobjekten", PhD thesis, Technical University Ilmenau, Germany, 2007 (in German).

[4] Van Eijk, R., Kuijsters, A., Dijkstra, K. and IJsselsteijn, W., A., "Human Sensitivity to eye contact in 2D and 3D videoconferencing", Proceedings of the QoMEX 2010, No. 76.

[5] Mekuria, R., Sanna, M., Asioli, S., Izquierdo, E., Bulterman, D. C.A., and Cesar, P., "A 3D Tele-Immersion System Based on Live Captured Mesh Geometry" ACM Multimedia Systems Conference, Oslo, Norway, February 27 – March 1, 2013.

[6] Kunze, K. and Strohmeier, S., "Examining subjective evaluation methods used in multimedia quality of experience research", Proc. 4th International Workshop on Quality of Multimedia Experience, Yarra Valley, Australia, (QoMEX 2012).

[7] Kepplinger, S., "Subjective Quality Assessment of Free Viewpoint Video Objects", Adj. Proc. 9th European Interactive TV Conf. (EuroITV.2011), Doctoral Consortium, 47-49.

[8] Dix, A., Finlay, J., Abowd, G. and Beale, R., "Human-Computer Interaction", (3rd edition). Harlow: Pearson Education Limited, 2004.

[9] Weiss A. et al., "User Experience Evaluation with a Wizard of Oz Approach: Technical and Methodological Considerations", Humanoid Robots, 303-308, ACM press, 2009.

[10] Li, Y., Hong, J. I. and Landay, J. A., "Design Challenges and Principles for Wizard of Oz Testing of Location-Enhanced Applications" Pervasive computing, 6(2), 70-75, 2007.

[11] Strauß, P., Hoffmann, H. and Scherer, S. "Evaluation and User Acceptance of a Dialog System Using Wizard-of-Oz Recordings", Proc. 3rd IET Intern. Conf. on Intelligent Environments, Ulm, Germany, (IE'07), 2007, pp. 521-524.

[12] Petrik S., "Wizard of Oz Experiments on Speech Dialog Systems. Design and Realisation with a New Integrated Simulation Environment", Diploma thesis, Technical University Graz, Austria, 2004.

[13] Skype (allows free video calls). Available online at: http://about.skype.com/. Last seen: 4th April 2013.

[14] Recommendation ITU-R P.805: "Methods for objective and subjective assessment of quality", Recommendation ITU-R P.805, ITU Telecommunication Standardization Sector of ITU, 2007.

[15] LabVIEW is a graphical development software. Available online at: http://www.amc-systeme.de/. Last seen: 9th June 2013.