METHODS FOR FACILITATION OF WIZARD-OF-OZ STUDIES AND DATA ACQUISITION

Michael Geiger, Ralf Nieschulz, Martin Zobl, Robert Neuss, Manfred Lang

Institute for Human-Machine Communication
Technical University of Munich, D-80290 Munich, Germany

[geiger | nieschulz | zobl | lang]@ei.tum.de, neuss@usaneers.de

ABSTRACT

Usability engineering is a challenge in preparation, execution and analysis. In this paper we describe several facilitation methods for test execution and succeeding data analysis. On the one hand, the experimental manager is supported by speech control and a semi automated test flow. On the other hand, the appearing data is logged in a way that allows to replay the whole experiment authentically afterwards and eases data analysis substantially. The combination of the described methods leads to more objective and consistent data as well as an efficient test execution and data survey.

1. INTRODUCTION

Testing in complex, multimodal usability scenarios can easily overstrain the experimental manager. Particularly when observing visual input like gestures the manager needs to pay visual attention to the test subject's activities in order to assure an immediate and accurate system response. Distraction of the test manager through keyboard input or reading out tasks has to be prevented because this may have a disturbing influence on the study and lead to test results that often are incomparable. In consequence error-free and objective conclusions cannot be drawn. Usually, this makes it necessary to have several test managers who arrange the duties. Furthermore, a usability study often leads to a huge amount of data, which makes an evaluation difficult and error-prone. Besides, analyzing video material generally is time consuming and also open for mistakes. Moreover, filling in test accompanying questionnaires is often bothering the test subject. This often leads to inconsistent, inconsiderate or even incomplete answers. Profiting by experiences from earlier studies [Zob01] we developed these methods to ease usability studies.

2. METHODOLOGY

The described methods ease the transaction of usability studies applying the so called 'Wizard of Oz' (WoO) methodology. In a WoO study the experimental manager – the 'wizard' – telecontrols occurring events and also influences the system's behavior, while the test person is told to interact with an already implemented and functioning system. Main goal of such studies is to survey an application for user-friendliness concerning its human-machine interface. While the test subject is confronted with the human-machine interface, the test manager resides in a separated control room. He has visual connection to the subject from several points of view. They communicate via audio intercom. In our system the different components communicate via network (TCP/IP) and may run on different computers. An overview of the whole system architecture is shown in figure 1.

2.1. Semi automated task control

We use a semi automated task control to run the test sequence. In the 'task storage' all occurring user tasks are sequentially stored as pre recorded speech. The test manager only causes the system to provide the user with the respective task from the 'task storage' via audio output. This assures that every subject obtains exactly the same instructions under the same conditions.

2.2. Wizard Speech Recognition (WISPER)

The experimental manager uses a command language speech recognition we call WISPER (Wizard Speech Recognition) to control the semi automated test sequence (e.g.: "read out next task") or to directly interfere the test scenario (e.g.: initiating a phone call with the subject) by voice. Furthermore, the 'wizard' translates observed user actions into speech commands. These commands are then processed by the 'input simulator' which imitates real user input by sending respective commands to the application. System input can alternatively be done by the user directly via haptic console or real speech / gesture recognition.

From: Proc. HCI 2001, Lawrence Erlbaum Ass., NJ

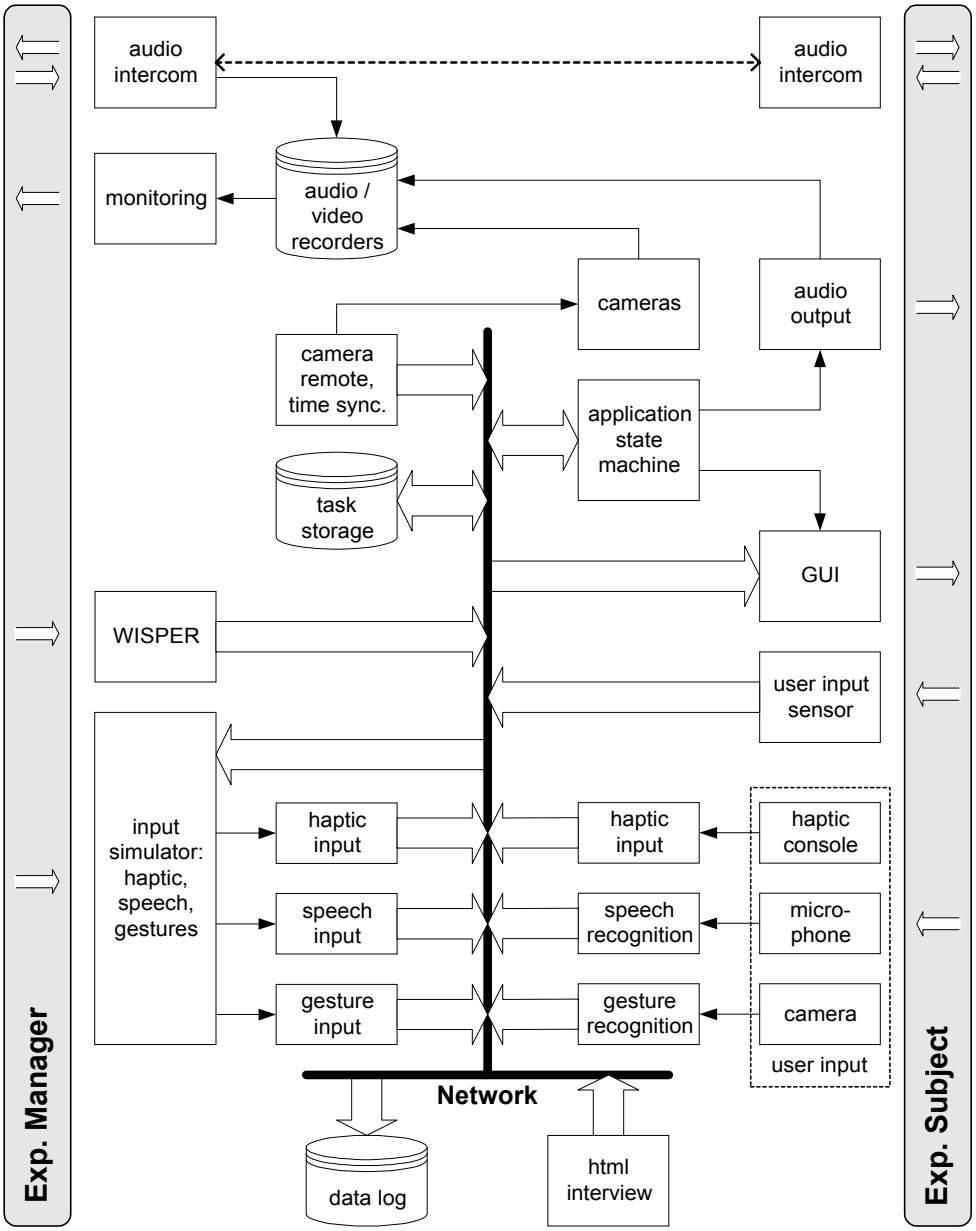

Figure 1: System overview

WISPER reduces the experimental manager's mental work to a minimum compared to haptical input and allows him to focus on the test subject's actions exclusively without averting his eyes from the observation monitor. From the subject's point of view, the response time of the system becomes very short. Generally, this is an important criterion for WoO studies.

2.3. Data recording

Throughout the whole test procedure all relevant data is automatically logged to hard disk with time stamp. It contains the status of the MMI and all actions initiated by the test manager as well as by the subject. To obtain objective information about the subject's reaction and processing time, a 'user input sensor' is applied. This can be a simple button, which has to be pressed by the subject permanently except while performing a task. This information is also stored in the data file. Later, this data can be easily evaluated by further applications like MATLAB (© The MathWorks, Inc.) or processed by another application reconstructing the complete test-flow authentically. As extensive information is automatically stored in the log file, time-consuming analysis of video recordings is considerably reduced.

2.4. Audio and Video Recording

The cameras show the subject from several angles. These video signals are mixed picture in picture with the GUI presented to the subject. Also the system time is noted on the video recording. As video material and data log are synchronized to each other, it is very easy to match certain data with the respective video sequence afterwards. Besides, all audio signals (intercom between 'wizard' and subject, audio output etc.) are recorded on the video tape.

2.5. Html Interview

In order to gather subjective impressions of the test person, it is necessary to interview him or her repeatedly during the experiment. We use html based questionnaires that are displayed to the subject overlaying the GUI. The experimental manager leads an interview by reading out the questions and filling in the html form according to the subjects answers. As a result, the subject is not bothered but stimulated to answer considerately. Furthermore, misinterpretations of questions can be solved immediately by call backs, which must not influence the subject's opinion. Before saving, the completeness of the questionnaire is checked automatically by a cgi-script. The data format is optimized for being processed by other applications later on.

Figure 2: The experimental manager in the control room

3. RESULTS & CONCLUSION

The combination of the methods described above was used so far in two WoO studies [Gei01][Nie01] and proved as a stable, efficient and consistent procedure. Due to the modular structure of the system it can be easily expanded. The automated data acquisition leads to a noticeable reduction of analyzing time and to more consistent data. The possibility to reproduce the whole test sequence allows us to find errors in the conceptual formulation during preliminary investigation as well as bugs in the surveyed application. Finally, even a complex usability scenario can smoothly be managed by only one person.

REFERENCES

[Gei01] Geiger, M., Zobl, M., Bengler, K., Lang, M. (2001). Intermodal Differences in Distraction Effects While Controlling Automotive User Interfaces. *Proceedings HCII 2001 (New Orleans, Louisiana, USA), (this conference*).

[Nie01] Nieschulz, R., Geiger, M., Bengler, K., Lang, M. (2001). An Automatic, Adaptive Help System to Support Gestural Operation of an Automotive MMI. *Poster Proceedings HCII 2001 (New Orleans, Louisiana, USA), (this conference*).

[Zob01] Zobl, M., Geiger, M., Bengler, K., Lang, M. (2001). A Usability Study on Hand Gesture Controlled Operation of In-Car Devices. *Proceedings HCII 2001 (New Orleans, Louisiana, USA), (this conference*).