A tool for Wizard of Oz studies of multimodal mobile systems

Carmelo Ardito, Paolo Buono, Maria Francesca Costabile, Rosa Lanzilotti, Antonio Piccinno
Dipartimento di Informatica, Università di Bari, Italy
{ardito, buono, costabile, lanzilotti, piccinno}@di.uniba.it

***Abstract* — The MuMoWOz (MultiModal Wizard of Oz) tool presented in this paper permits to conduct Wizard of Oz user studies. The significant features of MuMoWOz are the possibility to easily adapt it to new simulation scenarios and an architecture that allows HCI researchers to perform studies in mobile multimodal settings. The simulation of two scenarios related to cultural heritage domain is reported.**

***Keywords* — Wizard of Oz, Interaction Design, User Study**

I. Introduction

User-centered design has been proven as a key factor for leading towards the development of successful interactive systems [1, 2, 3]. In the system development process, it implies that final users are involved from the very beginning of the planning stage. Early involvement of users has the potential for preventing serious mistakes when designing innovative systems. The basic principles of user-centered design are: 1) analyze users and task; 2) design and implement the system iteratively through prototypes of increasing complexity; 3) evaluate design choices and prototypes with users.

During the early design stage it is compelling for the design team to understand users requirements, and to test the design choices quickly and informally, so providing feedback to the design activities. The Wizard of Oz (WOz) technique [4] is very useful in situations where the development of a system is expensive and it is hard to know beforehand how users will behave. A simulation of the final system is quickly built to be managed by the wizard, and end-users have the feeling they are interacting with the actual system.

The MuMoWOz (MultiModal Wizard of Oz) tool presented in this paper has been developed in order to test simulation of actual prototypes of multimodal mobile systems, either in laboratory settings or in delimited environments, using the WOz technique. Respect to other tool, the main contribution of MuMoWOz is its generality: it is possible to simulate the user interaction in various scenarios of different domains by only setting the appropriate multimedia content.

Section II discusses some important issues, which arise evaluating mobile systems. In section III the MuMoWOz tool is described. The simulation of two systems designed for the cultural heritage domain is reported in section IV. The Related Work, in Section V, reports other WOz tools developed to test systems in various domains. Section VI concludes the paper.

II. Evaluating mobile systems

In the field of mobile systems the WOz technique is especially useful when researchers wish to experiment with new interaction techniques to understand how to possibly overcome the devices constraints or the available technology is not tailored to the design plans.

Mobile computing presents new challenges in terms of evaluation techniques. While task-centric evaluation approaches may be well applicable to the desktop computing paradigm, where tasks are usually structured and almost predictable, they are not directly applicable to the often unpredictable and unstable mobile settings.

Field-based evaluations seem an indispensable approach for evaluating the usability of mobile systems. Yet, evaluating usability in the field is not easy [5, 6]. It is far from trivial to apply established evaluation techniques, such as observation and think-aloud. Furthermore, field evaluations complicate data collection and limit control, since users are moving physically in an environment with a number of unknown variables potentially affecting the set-up [7, 8]. These difficulties are significantly reduced conducting laboratory-based usability tests, but some factors and issues that occur in the field are not addressed in laboratory settings.

Our point of view is that laboratory and field evaluations of a mobile system are complementary. Laboratory tests are cheaper than field tests and can be used in the early phases of the interaction design of a mobile system, because they provide a controlled environment within which researchers can isolate the effect of different variables on the test subjects. Once system prototypes have been positively evaluated in laboratory and the system is almost in a final version, we recommend to test it in the field. Permitting to simulate running prototypes, MuMoWOz improves the quality of the laboratory tests.

Partial support for this research was provided by Italian MIUR (grant "CHAT").
Responsible author: Carmelo Ardito, ardito@di.uniba.it, tel. +390805443289, fax: +390805443300

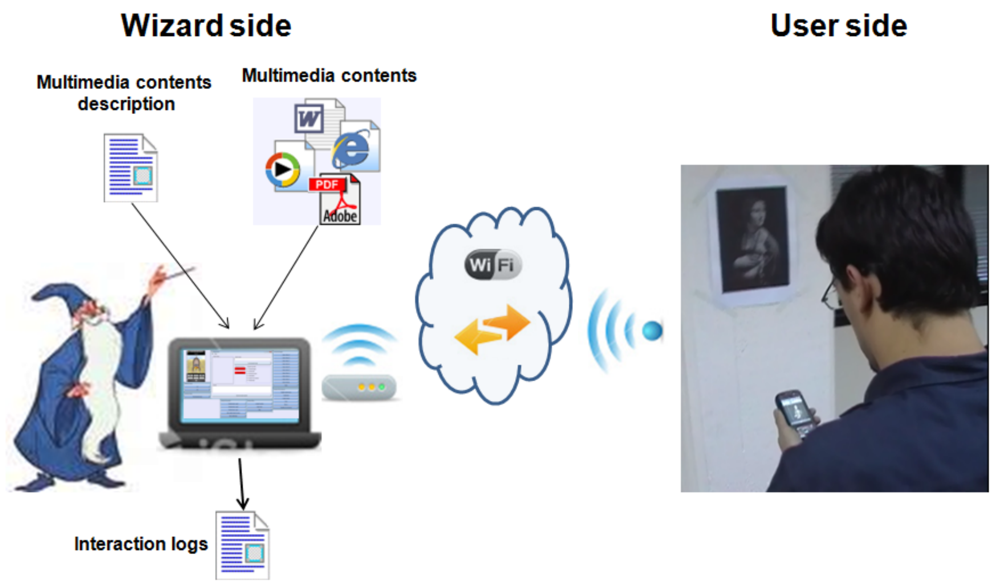

Fig. 1. MuMoWOz architecture.

III. MUMOWOZ: THE MULTIMODAL WIZARD OF OZ

A. System description

MuMoWOz is based on a distributed software architecture in which two macro-components, the server and the client, are connected by a wireless network. The server, installed on a notebook or a desktop computer, allows the wizard to send and manage the running of multimedia content at the client installed on a mobile device or another desktop computer and used by a participant to the study (see Fig. 1).

The client captures the interaction on the device and send it to the server. It is possible to keep track of the users interaction and to listen to them during the experiment, and thus to easily simulate the recognition of inputs, including those of multimodal type. In fact, the recognition of the particular input is not carried out by the system but by the human wizard, which interprets the user input and sends the required multimedia material that is presented on the client. This material, consisting of digital resources of various natures (HTML documents, video files, images, audio files, etc.), has to be prepared before carrying out the simulation and must be able to satisfy any possible user request in the scenario in question. For example, simulating a multimodal museum guide, if the user is in front of a paint s/he will probably ask (by voice, sketching or selecting interface widgets) more information about the paint or its author. The wizard reacts quickly pushing the button that visualizes on the client display the suitable document or video. If the user can not be satisfied, the wizard can play a predefined audio message (e.g. "this information is not available") or can write a short message, which is audio synthesized and played on the client.

The multimedia contents are described in a specific format. This description is inserted in a file needed by the application to keep track of the available material and to generate the Wizard's interface. Files reproduction is carried out by the applications installed on the client device: for example, a pdf (Portable Document Format) file is displayed by the installed pdf reading program, while a video can be played by the predefined player.

B. Wizard's interface

The wizard conducts the experiment using the Wizard's interface of the server component, which is installed on a desktop computer or a notebook. The Wizard's interface permits to give a rapid response to user requests so as to satisfy the simulation goals. The interface contains a main window and some panels with several buttons (see Fig. 2).

The *main window* allows the human wizard to:

- Control the interaction between the user and the client device. The wizard receives text feedback (which is reported by means of text messages printed in the large white area named "Sessione" or Session, in Italian) about the commands given by the user. At the end of the experimental session, all the collected data about the experimental session can be saved in a file, for use in an possible post-experiment phase.
- Digit sentences in the "Sintesi Vocale" text area (Speech Synthesis in Italian), which the system translates into vocal messages other than those available and predefined before the simulation was run. This makes the tool more flexible because it permits to face unexpected situations.

Observe, on the left side of Fig. 2, what is currently visualized on the client display. For example, the wizard can see that the portrait of the "Lady with the ermine" and two buttons to go back or forward are currently shown. Furthermore, in the "Sketch Viewer" area, the sketches drawn by the user on the device with the stylus are represented. In this way the user gestures can be interpreted by the wizard. In the example shown in Fig. 2, having drawn a "+", the user has asked for more information about the painting s/he is currently examining.

- Close the processes running on the client side or any open menus of the client interface through the two red buttons.

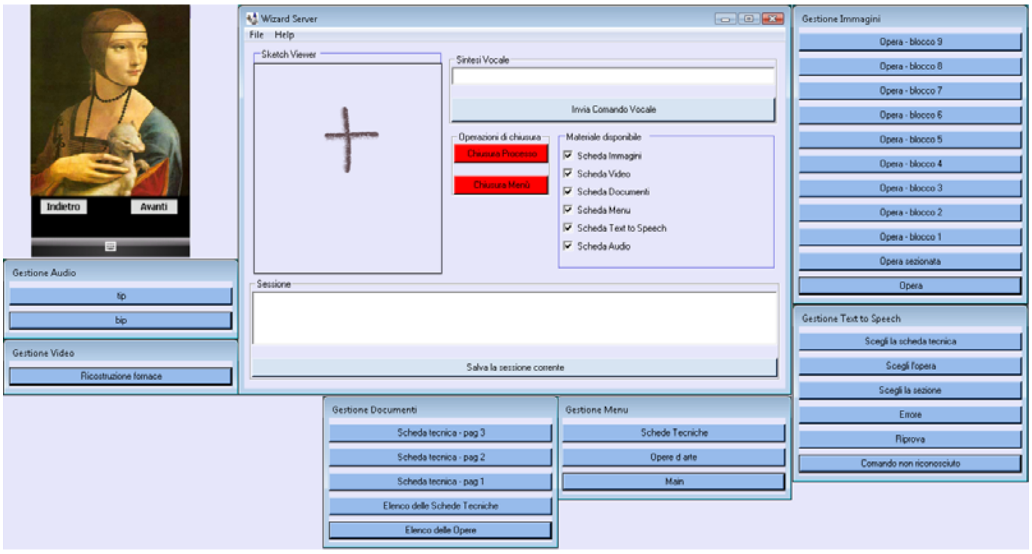

Fig. 2. The Wizard's interface.

- Personalize the aspect of the Wizard's interface by selecting the checkbox to display only the panels needed for the current simulation.

The *panels* contain the buttons for selecting the multimedia contents. The wizard can manage the reproduction of the multimedia material on the client device by means of such buttons. So, after interpreting the user input, the wizard can rapidly send the right content responding to the user request. The contents are classified in some categories: video, images, documents, text to speech, menu, audio. In the example shown in Fig. 2, the contents, belonging to six categories, are grouped in the six panels displayed. This facilitates the wizard's search for the content without getting lost in useless searches that would spoil the efficacy of the simulation.

IV. Simulating Scenarios using MuMoWOz

MuMoWOz has been designed and developed within the CHAT ("Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies") research project, which aims at developing a software infrastructure that provides services accessible through thin clients such as cellular phones or PDAs. Such services must be: a) adaptable to personal preferences of the user, with focus on the choice of interaction modalities; and b) adaptive to the physical-virtual context of the human actor carrying the device.

In the CHAT project MuMoWOz has been used to simulate two multimodal scenarios related to cultural heritage domain. These studies are reported in the following subsections. In both experiments, the wizard was seated away from the simulation area and has been supervised by an HCI expert. MuMoWOz has been installed on a notebook connected via a WiFi network to the mobile device carried by the user.

A. A game to learn archaeological park history

MuMoWOz has been adopted in the design of Explore!, an m-learning system that supports middle school students during a visit to an archaeological park. It adopts a learning technique called *excursion-game*, whose aim is to help students to acquire historical notions while playing a game on a cell phone. The study was performed in a university laboratory, involving four students that have already once played the traditional game during a school visit to the Egnathia archaeological park. Based on photos of the real site posted on the walls, the students were able to recall the site they had visited, thus simulating their presence in it. Students interacted with a prototype of Explore!, and the wizard, through MuMOWOz interface, sent all the multimedia material necessary for carrying out the game (i.e., text and audio messages, videos and images representing 3D models of interesting places of the park).

MuMoWOz allowed us to identify some interaction problems. During the evaluation, the students played the game in two different ways: with and without the audio modality. In the audio modality, the system beeps to capture students' attention, then gives spoken messages to inform users about application actions, i.e. the start of a new challenge or the transition from one phase to the next. After playing the game, the students were interviewed. They interacted pretty easily with the system and said they greatly appreciated the both electronic versions of the game, particularly the audio version of the system.

On the base of the results of this study and the students requests, we developed a running prototype of Explore! by inserting visual and sound messages that warn the user about what is happening, the map of the park, and giving new modalities to interact with the 3D models.

B. A context-aware multimodal museum guide

In this scenario a user visiting a museum can interact with a context-aware multimodal application installed on a PDA. The (simulated) system is able to recognize the position of the user in the museum thus presenting information related to the object in the proximity. It can also interpret user multimodal inputs (vocal input and sketch). The goal of the study was to observe users behavior interacting with such multimodal and context-aware museum guide.

The experiment has been performed in a university laboratory, which represented a room in the museum. Some paintings were posted on the walls. Every painting has been labeled with its name, author and period. A microphone has permitted to listen to the user's voice. Ten participants, selected among the students of the HCI course of the Laurea degree in Computer Science, have been involved.

The multimodal museum guide has been simulated using images and audios. When the user comes near to a painting, its image is automatically visualized on the display. The user can vocally ask for magnify to full-screen. S/He can obtain details about the painting currently in front of her/him by audio, text or video. The user can examine the other paintings of the same artist or those which have some relation. At the end of the experiment, participants answered a questionnaire regarding difficulties performing the tasks, their opinions about the interaction modalities, and their satisfaction using the system.

Various findings derived from the study; the most important was that the participants prefer to interact using vocal input rather than sketch, because they find stressful to remember the command connected to every symbol they can draw on the screen.

V. RELATED WORK

The WOz technique is not new [4]. It has been used in several situations. For example, to test natural language dialogue systems [9] and multimodal systems [10]. In [11], it has been applied to speech-based, multimodal ubiquitous computing applications, by substituting the speech recognition module of the system with a control application used manually by the wizard. WOz approach has been also employed in the design of computer vision based action games controlled with body movements [12].

In literature there are other examples of WOz tools. Singh et al. use immersive environments, that needs special equipment in order to work [13]. Another interesting WOz tool is DART a complex system that supports the iterative design process. Each research group has developed its own WOz tool to perform their specific studies; in most cases, a new tool must be designed and developed to execute an experiment with slight differences from a previous one.

In a Wizard of Oz study, subjects are intended to believe that they are using a computer system, while a person behind the scene plays the role of the program. Small typing errors, slow responses, no responses, or wrong responses can break this illusion. The wizard has to know every possible step the user might take next to be able to anticipate on these possible actions. To reduce this kind of problems, MuMoWOz requires that, before executing the test, the researchers must design in detail the simulation scenario. In this way, they can foresee most of the users actions and prepare in advance the multimedia content, which is represented on the Wizard's interface by "quick-to-select" buttons arranged in specific panels.

VI. CONCLUSION

System development is a very costly process, from both the time and the economic standpoints. This problem is even more evident in the field of mobile devices, especially in situations where the aim is to experiment new interaction techniques, or to take into account the surrounding environment and so requiring the integration of sensors and actuators. Being quickly adaptable to different situations, MuMoWOz can be used for WOz studies to check if the available technology is tailored to the design plans and to make decisions on alternative design choices before investing heavy resources.

ACKNOWLEDGMENT

We thank students M. De Bartolo, and A. Giannelli for their contribution to implementing the MuMoWOZ tool.

REFERENCES

[1] A. Dix, J. Finlay, G. Abowd and R. Beale, *Human Computer Interaction*, Prentice Hall, 1998.

[2] J. Preece, Y. Rogers, H. Sharp, D. Beyon, S. Holland and T. Carey, *Human-Computer Interaction*, Addison Wesley, 1994.

[3] J. Rubin, *Handbook of Usability Testing*, John Wiley& Sons, 1994.

[4] N. Fraser and N. Gilbert, Simulating Speech Systems. *Computer Speech and Language*, 5, 1991. pp. 81–99.

[5] S. Brewster, Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing*, 6, 2002. pp. 188-205.

[6] C. Nielsen, Testing in the Field. Proc. *APCHI'98*, IEEE Computer Society. pp.285-290. 1998.

[7] P. Johnson, Usability and Mobility: Interactions on the move. Proc. of the *First Workshop on Human-Computer Interaction with Mobile Devices*, Glasgow. 1998. Available on-line.

[8] H. Petrie, V. Johnson, S. Furner and T. Strothotte, Design Lifecycles and Wearable Computers for Users with Disabilities. Proc. of the *First Workshop on Human-Computer Interaction with Mobile Devices*, Glasgow, UK. 1998. Available on-line

[9] M. Johnsen, T. Svendsen, T. Amble, T. Holter and E. Harborg, TABOR – A Norwegian Spoken Dialogue System for Bus Travel Information. Proc. *ICSLP 2000*, ISCA Archive. Vol.3, 2000. pp. 1049-1052.

[10] J. Gustafson, L. Bell, J. Beskow, J. Boye, R. Carlson, J. Edlund, B. Granstrom, D. House and M. Wiren, AdApt - A Multimodal Conversational Dialogue System in an Apartment Domain. Proc. *ICSLP 2000*, ISCA Archive. Vol.2, 2000. pp. 134-137.

[11] K. Mäkelä, E-P. Salonen, M. Turunen, J. Hakulinen and R. Raisamo, Conducting a Wizard of Oz Experiment on a Ubiquitous Computing System Doorman. Proc. *International Workshop on Information Presentation and Natural Multimodal Dialogue*, Verona, Italy, December 14-15, 2001. pp. 115-119.

[12] J. Höysniemi, P. Hämäläinen and L. Turkki, Wizard of Oz prototyping of computer vision based action games for children, Proc. *IDC2004*, ACM Press, New York. pp.27-34.

[13] P. Singh, H. N. Ha, Z. Kuang, P. Olivier, C. Kray, P. Blythe, and P. James. Immersive video as a rapid prototyping and evaluation tool for mobile and ambient applications. In MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, pages 264{264, New York, NY, USA, 2006. ACM Press.

[14] S. Dow, B. Macintyre, J. Lee, C. Oezbek, J. Bolter, and M. Gandy, "Wizard of Oz Support throughout an Iterative Design Process," Pervasive Computing, IEEE, vol. 4, 2005, pp. 26, 18.