OBSERVING USERS IN MULTIMODAL INTERACTION

Daniel Salber

Laboratoire de Génie Informatique—IMAG
BP 53, 38041 Grenoble Cedex 9, FRANCE
Phone: +33 76 51 44 40, Fax: +33 76 44 66 75, E-mail: daniel.salber@imag.fr

ABSTRACT

My doctoral research focuses on the usability and usage of new computer technology such as interactive systems that support the combination different input media such as voice, gesture and video. I have contributed to the study of these systems in four complementary ways: the MSM framework helps classifying and reasoning about current and future "interactionally-rich" systems; I have also used formal methods to specify multimodal interaction; I have designed and developed the NEIMO multimodal Wizard of Oz platform, a tool for observing users using multimodal interaction and assessing usability issues; I am now working on computer-mediated communication and collaboration through audio and video.

KEYWORDS: Multimodal Interaction, Evaluation, Usability, Formal Methods, Observation of Users, Video, Groupware, CSCW, MediaSpace.

INTRODUCTION

My doctoral research focuses on the usability and usage of new computer technology such as interactive systems that support the combination different input media such as voice, gesture and video. There is a high potential for systems allowing the use of combined input and output media but our knowledge for designing, building, and evaluating such systems is still primitive. My primary goal is to clarify and structure such knowledge from the system perspective. I have contributed to this objective in four complementary ways: first, I have participated to the definition of the MSM framework to help reasoning about current and future systems; second, I have used formal method to specify multimodal interaction; then, I have designed and developed NEIMO, a generic computerized Wizard of Oz platform for assessing and observing the usability and usage of multimodal settings; I am currently investigating human-to-human communication and collaboration through audio and video connections.

THE MULTI-SENSORI-MOTOR FRAMEWORK

The underlying motivation for the Multi-Sensory-Motor framework (MSM) is to provide designers with a structured problem space to classify and reason about interactionally-rich systems. MSM is comprised of 6 dimensions: two of them deal with the notion of communication channel: the number and direction of the channels that a particular system supports. The other four dimensions are used to characterize the degree of built-in "cognitive sophistication" of the system (i.e., the interpretation and rendering functions of the system). They include the notions of levels of abstraction, context, fusion/fission, and granularity of concurrency. In addition to set up a clear distinction between the system properties of multimediality and multimodality, MSM has made explicit software design pecularities of interactive systems such as the fusion of user inputs along different media. One interesting thing is that fusion (and fission) do occur in current graphical user interfaces as they do in the "put that there" paradigm but this phenomena was not accounted for in an explicit way. MSM helped us to integrate such general phenomena into a uniform conceptual software architecture model [1].

EXPERIENCE WITH THE USER ACTION NOTATION

Formal methods can be used either as a specification tool for describing the user interface before it is developed, or as an analysis tool to assess properties of the system. This second approach is very relevant to usability issues: it is often the case that characteristics that can be assessed are also usability properties. Unfortunately, currently available formalisms do not support the description of multimodal user interfaces. From modelling MATIS, a multimodal system developed in our research group [7], I have extended UAN [4] in order to express multimodality. In addition, I have devised a preliminary set of heuristic rules to formalize usability problems based on an extended UAN specification. Such analysis allows the designer to track a particular class of usability problems during the specification phase. Empirical observations of users such as those supported by NEIMO, allow the detection of possibly different classes of design errors.

NEIMO: A WIZARD OF OZ PLATFORM FOR MULTIMODAL INTERACTION

NEIMO is a generic, multi-wizard, multi-subject, computerized platform that supports Wizard of Oz (WOz) experiments for the observation and analysis of multimodal interaction [8]. WOz systems have been developed on a case-per-case basis and support the observation of one modality only. Similarly, automated analysis tools are limited in scope and are rarely integrated into the WOz platform from the start. Although there has been a few attempts to apply the WOz technique to specific multimodal systems [2, 6], there has been no attempt to produce a generic, reusable WOz platform such as NEIMO that would make possible the observation and analysis of multimodal interaction. During an evaluation session, NEIMO records digital data at different levels of abstraction (i.e., from keystroke/event level up to commands and domain concepts level). This material can then be used by evaluation experts to assess the usability of the system. A collaboration with human factor experts from the Hewlett-Packard usability lab in Grenoble pointed out the major part played by the use of video in the evaluation process and led us to enhance the video observation capabilities of NEIMO. The platform is currently being used by psychologists to identify mental strategies used in linearization tasks [5]. It is also involved in testing the usability of an innovative industrial telecommunication product and for exploring ways of extending the current design with multimodal capabilities.

VIDEOPORT: A MEDIASPACE EXPERIMENT

The multi-subject and multi-wizard capabilities of NEIMO have also led me to investigate the collaboration process in a work group: the human wizards need to collaborate to perform the services simulated during the WOz session. Visits at Rank Xerox EuroPARC in Cambridge, UK, have demonstrated the interest of MediaSpaces, both to support a collaborative task such as the wizards' and as a support for informal communication and collaboration. I developed VideoPort, a small-scale MediaSpace for our research team, that will also be used to support the wizards' task during WOz evaluation sessions. This MediaSpace has some original features: it is entirely digital, and provides advanced privacy protection mechanisms. To investigate further the usage of video in computer-mediated communication and collaboration, I recently spent three months at the MRC-Applied Psychology Unit in Cambridge, UK with Phil Barnard. During my stay, I set up and ran an experiment to assess the use of different camera positions for the understanding of deixis in an audio-video environment. This experience was also an opportunity to learn about psychology subject-testing protocols. This knowledge will be very useful for our usability tests with NEIMO.

PERSPECTIVES

From now on, I plan to continue working on the use of video and studying the potentials of video for communication and collaboration in a work group. I plan to extend our VideoPort system, first with multimodal input capabilities, such as voice commands and then with face recognition capabilities, based on innovative work done by computer vision researchers in Grenoble [3].

REFERENCES

1. J. Coutaz, L. Nigay, D. Salber, J. Caelen: *The MSM Framework: A Design Space for Multi-Sensory-Motor Systems*, InterCHI'93 Workshop on Multimedia & Multimodal Systems, Amsterdam, The Netherlands.
2. N. Dahlbäck, A. Jönsson and L. Ahrenberg: *Wizard of Oz studieswhy and how*, 3rd Conference on Applied Natural Language Processing, Trento, Italy, 1992.
3. P. De Marconnay, J. Crowley, D. Salber: *Visual Interpretation of Faces in the NEIMO Multimodal Test-Bed*, IJCAI'93 Conference, Chambéry, France
4. H.R. Hartson, P.D. Gray: *Temporal Aspects of Tasks in the User Action Notation,* Human-Computer Interaction, 1992, Vol.7, pp. 1-45.
5. W.J.M. Levelt: *Linearisation in describing spatial networks*. in S. Peters & E. Saarinen (Eds.), Processes, beliefs, and questions, Dordrecht, NL.
6. D. Maulsby, S. Greenberg, R. Mander: *Prototyping an Intelligent Agent through Wizard of Oz*, InterCHI'93 Proceedings.
7. L. Nigay, J. Coutaz, D. Salber: *MATIS: A Multimodal Airline Travel Information System*, AMODEUS 2 Project Document, RP1, WP 10
8. D. Salber and J. Coutaz: *A Wizard of Oz Platform for the Study of Multimodal Systems*, InterCHI'93 Conference Adjunct Proceedings, Amsterdam, The Netherlands.