INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team MErLIn

Methods for Interactive Software Ergonomics

Nancy - Grand Est

THEME COG

Table of contents

1. Team

Faculty Member

Noëlle Carbonell [Team Leader, Professor, Henri Poincaré University, LORIA, HdR]

Technical Staff

Jérôme Simonin [Doctor-Engineer, INRIA-Nancy-Grand-Est, from November 1st, 2007]

Post-Doctoral Fellow

Andrei Lintu [INRIA-Nancy-Grand-Est, from December 31, 2007]

Administrative Assistant

Laurence Benini [TR, INRIA, part-time]

Other

Olivier Christmann [Lecturer, Henri Poincaré University, LORIA, from September 1st, 2007 until August 31, 2008; Lecturer, ENSEM-INPL, LORIA, from September 1st, 2008]
Daniel Gepner [Lecturer, ENSEM-INPL, LORIA, from November 1st, 2007]
Gaétan Pruvost [5th year Student in Computer Science, Supelec Metz and Master Informatique des Universités de Nancy, Spécialité Recherche PRIM, LORIA, from February 1st until August 31, 2008]
Anne-Sophie Tranchet [3rd year Student in Computer Science, Henri Poincaré University, LORIA, from March 1st until June 30, 2008]

2. Overall Objectives

2.1. Introduction

Keywords: *3D gestures*, *3D interactive visualizations*, *adaptive user interfaces*, *ergonomic quality of interactive software*, *gaze*, *gaze-contingent displays*, *multimodal interaction*, *online help*, *speech*, *user assistance*, *user interface design and evaluation methods*, *user modeling*, *user testing*, *visual search*.

The goal of the MErLIn team is to contribute to the improvement of the Ergonomic Quality of Interactive Software. Two sub-goals contribute to this general goal:

- The study, using empirical or experimental approaches, of users' interactions with software implementing novel concepts in emerging application domains, so as to increase available knowledge of users' perceptual and cognitive behaviors in realistic contexts of use, and to evolve effective design guidelines from the results of these studies.
- The design, implementation and evaluation of new Human-Computer Interaction (HCI) paradigms that take advantage of recent scientific advances in user modeling and new media processing. The main aim is to explore emerging HCI design spaces with a view to proposing new interaction paradigms that are easy to implement reliably and that satisfy the needs, expectations and capabilities of future users.

The MErLIn team contributes to scientific advances in the areas of user assistance, multimodal interaction, visual search in picture banks, and computer accessibility (especially in the context of 'Ambient Assisted Living'). Recent scientific publications of the team include papers accepted at major Conferences or Journals in these four areas: Adaptive Hypermedia (AH08), ACM International Conference on Multimodal Interfaces (ICMI08), ACM Working Conference on Advanced Visual Interfaces (AVI06) and Visual 2008 (Int. Conf. on Visual Information Systems), UAIS (Int. J. Universal Access in the Information Society, Springer), IJMUI (Int. J. Multimodal User Interfaces) . Main outcomes of the team's research activities include ergonomic recommendations, prototypes and software platforms for supporting verisimilar simulations of new HCI paradigms, thus making it possible to collect realistic, empirical or experimental, data on future end users' behaviors, subjective judgments and non verbal responses.

Currently, the MErLIn team investigates usability issues raised by "new" application domains and new paradigms or forms of interaction (new technology often raises new usability problems). The aim is to acquire novel ergonomic results on innovative software systems, and to further increase current knowledge on usability. Research efforts are currently focusing on the design and assessment of:

- Enhancements to online help intended for novice users in the general public: (i) adaptive help (i.e., dynamic adaptation of message content to the current knowledge and skills of the user), (ii) multimodal (speech and graphics) presentation of help information, (iii) embodiment of the help system using a humanoid 3D talking-head (emerging paradigm of affective interaction).
- Novel modalities or forms of multimodality for interacting with virtual reality environments: enhanced gaze-contingent displays, gaze (as a pointing modality) combined with speech, 3D gestures.
- Interactive 3D visualizations meant to facilitate visual search in large unstructured photo collections.
- Software environments with gaze-contingent display facilities (i.e., 3D animated scenes including embodied conversational agents) intended for improving the social communication skills of autistic teenagers (e.g., Asperger syndrome).

2.2. Highlights of the year

This year, it is worth noticing:

- The beginning of a multi-disciplinary research project supported by the 'Fondation de France' on autism; partners include computer scientists (at LIMSI), psychologists (at UMR 7593, Centre Emotion), and clinician physicians (at the 'Centre Hospitalier' in Chartres).
- The defense of Olivier Christmann's PhD thesis on visual search in large unstructured picture collections.

3. Scientific Foundations

3.1. Definitions

The scientific domains characterizing the activities of the MErLIn team are Human-Computer Interaction and Software Ergonomics.

Software Ergonomics is a science that contributes to the knowledge necessary to software design and, more generally, to computer-based environments, with the overall perspective of human security and well-being, but also with the perspective of effectiveness, efficiency and productivity, for instance by facilitating users' tasks, limiting learning time, reducing errors and the cost of errors. Software Ergonomics focuses on the improvement of human-computer interactions mainly in terms of cognition, as the main human activity involved with software interactions is mental. However, as novel interaction techniques (e.g., multimodality) and novel environments (e.g., virtual reality) arise, some aspects of physiological ergonomics have to be considered as well.

Ergonomic Quality of Interactive Software covers usability (or ease of use) in the widest sense, that is, the extent to which users can easily reach their interaction goals (presentation and dialog aspects), and utility, that is, the extent to which users can reach their task goals (functional aspects of the interaction such as functions, objects, data, etc.).

Human-Computer Interaction (HCI)[1] is also at the center of the team's research activities. Human-Computer Interaction can be defined as the set of hardware, software, human and environmental elements that influence the efficiency of systems and products, both from a technological and a human point of view.

[1] Definition from AFIHM (Association Francophone d'Interaction Homme-Machine) (http://www.afihm.org)

3.2. Approach

The MErLIn team uses methods from Ergonomics and Software Engineering, with a strong background and experience in experimental approaches and methods in the sense of experimental sciences, that is, with hypotheses testing and proving.
The team contributes to user and interaction modeling from experimental testing in the laboratory or field simulations, using performance data, analysis of verbal or multimodal protocols, analysis of subjective judgments and non verbal responses. Outcomes of modeling activities include insights into users' perceptual and cognitive processes, and computer models capable of predicting and taking into account users' behaviors, capabilities and expectations in complex interaction situations. The appropriateness and accuracy of such models compared to reality is determined through ergonomic evaluations.
Research usually starts from the observation of real tasks in a specific field of activity, sometimes in parallel with particular practical problems to be solved. Data gathering is based on activity and interaction analyzes, using interaction logs and video recordings.

3.3. Focus

MErLIn's research activities, which aim at integrating ergonomics approaches within the computer system life cycle through sets of design recommendations, methods and software support tools, focus on software intended for users who are not computer specialists. This user population is the major target of current software developments; it includes the general public (e.g., interactive booths, mobile systems, etc.) and professional experts in various domains (e.g., biology, history). The focus is not only on standard desktop interactions, but also on new contexts of computer use, such as: consumer products (e.g., electronic commerce), information retrieval (e.g., tourism), mobility, etc. Special attention is paid on accessibility which promotes increased effectiveness, efficiency, and satisfaction for people with a wide variety of capabilities and needs.

4. Application Domains

4.1. Application Domains

This year, the main application domains have been:

- photo browsers for the general public: design and ergonomic evaluation of 3D visualizations of large unstructured collections of pictures and interaction metaphors that facilitate visual search and improve its efficiency;
- online help for the use of software intended for users who are not computer specialists: assessment of the contribution of Embodied Conversational Agents (ECAs) to the efficiency and acceptance of online help, evaluation of the efficiency and acceptance of adaptive online help;
- multiscale displays of large and complex scenes (e.g., Google earth): enhancements to gaze-contingent displays.

Specific scientific results are presented in the next section.

5. New Results

5.1. Introduction

The team's scientific progress and contributions during 2008 have been grouped into three main topics: enhancements to online help for the general public, interactive 3D visualizations, multimodal interaction (gaze and 3D gestures).

5.2. Enhancements to online help for novice users in the general public

Participants: J. Simonin, A.-S. Tranchet, N. Carbonell.

To improve the effectiveness of online help, we are currently investigating two research directions: adaptivity and multimodality.
User modelling is an active, fast developing research area. According to [27], recent scientific advances make it possible to consider the implementation of effective adaptive user interfaces, that is, interfaces capable of adapting their behaviours to the evolution of the current user's profile, namely their competences, interests, preferences and/or goals. However, specific usability issues are yet to be investigated; see [8]. To react predictably is a major usability requirement for interactive software, according to Shneiderman. Hence, how will users, especially novice users, react to, and accept, user interfaces whose behaviors evolve autonomously during interaction? What amount of control over the interface evolution should users be given?
As for multimodality, we have demonstrated that multimodal contextual help combining oral messages with graphics will be well accepted by novice users; it will also stimulate help consultation, and improve its effectiveness [2], mainly because oral help messages do not disrupt novices' interactions with new software, contrary to textual messages. Embodying online help systems and endowing them with speech capabilities may further increase their effectiveness. Using an Embodied Conversational Agent (ECA) for assisting novice users, has the potential to enhance their motivation to consult online help systems more frequently, hence to facilitate and improve learning of how to operate new software.
This year, efforts have been focused on the analysis of multimodal protocols from two empirical studies: one was meant to assess the ergonomic quality of adaptive online help (9 participants), and the other aimed at identifying the actual contributions of ECAs to the efficiency and usability of help systems intended for novice users in the general public (22 participants). An animation creation software (Flash) was used for both studies. The functionalities of the two help systems were partly simulated using the Wizard of Oz technique, the Wizards being assisted in their activities by a generic software platform which also recorded participants' interaction traces.
The adaptive help system could dynamically adapt the information content of text+graphics messages to the evolution of novice users' knowledge of Flash operation; it also anticipated novice users' information needs, based on the Wizard's detection of their goals and intentions from their interactions with Flash. Oral messages illustrated with Flash screen copies were implemented in the multimodal help system; they were activated on the user's initiative exclusively. In one condition, oral messages were spoken by an ECA, a female talking head developed by FT R&D; in the other condition the same multimodal messages without the ECA were proposed to participants.
Results of these studies which are fully detailed and discussed in Jérôme Simonin's PhD thesis [10] have been presented at two international conferences ([15] and [14] and one francophone workshop ([16]). Concerning adaptive and multimodal online help, comparative analyzes show that subjective judgments were more positive for the adaptive system than for the multimodal one, while learning Flash operation proved to be less effective with the adaptive system than with the multimodal one. As for the presence of the ECA, it was well accepted by all participants; analysis of interaction logs indicates that it increased help consultation sensibly compared to the condition where it was not present.

5.3. Interactive 3D visualizations of large unstructured picture collections

Participants: O. Christmann, N. Carbonell.

The design of interactive visualizations of large collections of pictures has not yet raised much interest in the HCI research community. This lack of interest seems to be based on the implicit assumption that visualization techniques designed for large digital or textual data sets can be used for visualizing large collections of pictures. For instance, [21] experimented tree-maps for displaying photographs taken at CHI 2001. As for designers of photo browsers, they usually display photo collections in the form of scrollable 2D arrays of zoomable thumbnails.

However, specific visualization techniques need to be developed for the presentation of collections of pictures, especially personal collections. User motivations for browsing picture collections are different from those of users who explore digital or textual data sets. Search for a visually familiar photograph or for an unfamiliar photograph matching a set of criteria are two of the most frequent activities that motivate browsing through a picture collection, especially a personal one. In addition, most people are reluctant to annotate personal photographs, and they do not take the time necessary for classifying them appropriately; they just save them in folders with general, vague names. Therefore, most personal photo collections are unstructured for lack of appropriate metadata and/or manual annotations, whereas most visualization techniques have been designed for hierarchically structured data sets.

For several years, the team's research in this area has been focusing on the design, implementation and ergonomic evaluation of 2D and 3D interactive visualizations of picture collections, using standard interaction devices and modalities (i.e. mouse or joystick)[2]. Recent research efforts have been mainly focused on navigation and search in 3D visualizations of large photo collections (a thousand items or so), while earlier efforts were centered on visual search in 2D displays of small collections (i.e., 30 items); see [4] and [7] to get an insight into this earlier work.

Entertainment and commercial Web-sites, information kiosks and public terminals tend to display an increasing number of pictures simultaneously: video and movie snapshots, CD sleeves, book covers, etc. Personal electronic archives and file directories are increasingly cluttered with unstructured collections of photographs, scanned drawings, videos. The only option for searching large sets of pictures offered to users by current software (e.g., ACDSee, PhotoSuite or ThumbsPlus) amounts to scrolling 2D arrays of zoomable thumbnails. We have designed and implemented two 3D metaphors for visualizing and browsing large collections of photographs (e.g., landscapes, portraits, complex objects). Both metaphors visualize a collection of pictures or multimedia documents in the form of a vertical 3D cylinder; namely, arrays of thumbnails are plastered either on the inside or on the outside of the lateral surface of the cylinder. One cylindrical representation (ie., the outside view) incites users to view it as a virtual 3D object that can be freely manipulated (manipulation metaphor). The other representation (i.e., the inside view) may induce them to feel as if they were surrounded by a virtual cylindrical wall. In fact, this second representation suggests two different interaction metaphors: users may feel as if they were moving in front of the picture "wall" (locomotion[3]), or they may have the impression that they are moving the wall around themselves (manipulation).

These metaphors have been compared with respect to their possible influence on interaction efficiency (i.e., task execution times, success and failure rates, etc.) and usability (especially user subjective satisfaction). Participants carried out two types of search tasks: search for a visually familiar photograph or search for an unfamiliar photograph matching a verbal description; see [5].

This year, we analyzed the data from an experimental study which aims at comparing these two 3D representations with standard 2D array displays. Comparisons have been focused on the contributions of the three types of visualizations to search efficiency and comfort, and picture location memorization; see [30], which claims that 3D visualizations of large information sets facilitate memorizing item location compared to 2D representations. Tasks and experimental protocol were similar to those implemented in the previous study. That is, 20 participants carried out two types of realistic visual search tasks: looking for a visually familiar picture, and searching for an unfamiliar picture matching predefined criteria specified in writing. Each collection included about 1000 photographs, 150 of which (or so) being displayed simultaneously. Available actions on the three representations comprise left-to-right and right-to-left rotations, forward/backward adjustments and zooming facilities. Participants' eye movements were recorded together with interaction logs, so as to increase the number of objective measurements available for assessing the comfort and efficiency of visual search in each representation. Eye tracking data may also provide useful qualitative information on the possible influence of each representation on participants' visual exploration strategies. A large majority of participants preferred

[2] Our participation in the Micromégas project on multiscale visualization of, and interaction with, sets of familiar/personal information was an opportunity to develop research in this area. Micromégas was a three year national project (July 2003 - July 2006) in collaboration with the In Situ team at INRIA-Futurs and LRI (Orsay), and the LPM Laboratory in Marseille (Yves Guiard); it was supported by the ACI 'Masses de données'.

[3] More precisely, they may feel as if they were turning round on themselves, or "walking" along the virtual wall.

one of the 3D views to the standard 2D view. However, participants' performances were similar with the three views, due to high inter-individual differences in visual search strategies.
Results are detailed and discussed in Olivier Christmann's PhD thesis in computer science [12]. Some of them have been presented at an international conference [13]. A full paper will be soon submitted to Elsevier Journal on HCI research, 'Interacting with Computers' (IwC).

5.4. Multimodal interaction : gaze and 3D gestures

Participants: D. Gepner, A. Lintu, G. Pruvost, J. Simonin, N. Carbonell.

5.4.1. Gaze as a pointing modality

For the last three years our research efforts on gaze interaction have been focused on the investigation of two complementary research directions: (i) implementation of efficient gaze-contingent displays, (ii) design and implementation of multimodal, command languages.
Our main working hypothesis is that gaze may prove to be more appropriate than mouse or standard devices for designating objects and positions on the screen in contexts where the user interacts with very large displays, such as electronic walls, reality centers and caves.
We are working on the implementation and usability assessment of gaze as a pointing modality in environments where commands/actions can be expressed using speech or, in some application contexts, gestures. Using gaze in combination with speech or gestures makes it possible to design multimodal languages with an expressive power equivalent to, or even greater than, direct manipulation. It also offers the possibility to solve the "Midas touch" problem [26] transparently. We are using a head-mounted eye tracker (ASL-501).
Prior to addressing research issues we had to develop various software tools for recording eye tracking data, visualizing and analyzing them automatically. In particular, we had to develop an algorithm for computing fixations in real time; this algorithm in C++ takes head movements into account. We have also developed a software platform that can "replay" recorded user interactions with standard software applications (Windows or Linux) and superimpose the user's fixations on the successive displays. This year's progress is presented in the next two paragraphs. Our scientific activities on gaze interaction are summed up in [6].

5.4.2. Gaze-contingent displays

Participants: A. Lintu, J. Simonin, A.-S. Tranchet, N. Carbonell.

By definition, the resolution of gaze-contingent displays varies according to the user's current point of gaze, information density being higher around the current point of gaze than elsewhere on the display. The size of the high resolution area includes the foveal and para-foveal visual fields. To be effective and acceptable, gaze-contingent display algorithms should change display resolution as fast as the natural pace of human visual exploration of scenes. Research on gaze-contingent displays is developing rapidly. See [24], pages 285-299, for a review of current research in this area. Potential application areas include: (i) interaction with complex remote visualizations or animations the reactivity of which is still insufficient, due to transmission technique limitations; (ii) interaction with virtual reality environments such as reality centers or caves where display evolutions (e.g., viewpoint changes) often entail time-consuming computations.
Last year, we designed and developed a gaze-contingent prototype (60 Hz gaze sampling rate). The prototype is operational and has been experimented on blurred images displayed on a standard 21" screen. However, the delay necessary to detect fixations is too high (about 90 ms) for ensuring smooth scan paths; users resent to be slowed down in their exploration of the progressively de-blurred image.

We are currently investigating several strategies for getting round this difficulty. In particular, we are considering guiding gaze during scene exploration, since predicting the landing position of the next fixation from the speed and direction of eye movements during (or at the end of) the current saccade yields unreliable results [25]. To the best of our knowledge, gaze guidance, a difficult but promising research direction, has only been explored by a few research groups; see, for instance, [23]. We have experimented with a prototype that generates stimuli in the peripheral visual field, and observed that gaze is often, but not always, attracted by these stimuli. To achieve robust guidance of gaze movements, we have refined the implementation of this strategy and designed an experimental study meant to assess its effectiveness, and determine which properties peripheral stimuli should possess in order to ensure efficient and robust control of gaze moves (e.g., stimulus distance from the current point of gaze, stimulus size, duration, blinking frequency) without the user being aware of their presence. The positions of the stimuli on the 60 pictures constituting the visual material for this experiment have been defined from scan paths recorded earlier: one subject with normal vision viewed the 60 pictures without any constraint. Peripheral stimuli were placed on the pictures in the same positions as selected fixations from this subject, so as to reproduce his scan path and, thus, obtain realistic pre-defined scan paths. Therefore, the current experiment will make it possible to assess the influence of such "natural" scan paths on the efficiency of gaze guidance. Methodology and design issues have been solved; the visual material has been carefully selected and prepared (homogenization of photograph resolution and stimulus positioning); software tools necessary for implementing the designed setup have been developed and tested; see [11]. We are currently running a pilot study to adjust parameters and experiment with higher sampling frequencies of eye tracking data (120 Hz and 240 Hz) in order to increase system reactivity. The full experimental study will be run in December 2008 and January 2009.
The scope of this research is a priori limited to realistic free scene exploration activities (color photographs). Observation of scan paths during other activities, such as visual search for a particular item, will be performed later, with a view to determining whether and to what extent scan paths, hence gaze guidance, may be influenced by the type of visual activity users are currently engaged in.
The team participates in a national research project supported by the ‘Fondation de France’ (2008-2010). Other partners are researchers at LIMSI and UMR 7593, and clinicians at the ‘Centre Hospitalier’ in Chartres. The objective is to improve social communication skills of autistic teenagers’ (Asperger syndrome) by using virtual reality environments. The expertise of the team in gaze-contingent displays and gaze guidance is being used to focus autistic teenagers’ visual attention on the facial expressions of animated agents (ECAs) engaged in short dialogs, so as to improve their interpretation of non verbal information, and help them pay attention to, and better "understand", the attitudes and emotions underlying interlocutors’ verbal exchanges, such as joy, anger or irony.

5.4.3. Multimodal speech- and gaze-based interaction

Participants: D. Gepner, N. Carbonell.

Spoken natural language may appeal to users in the general public, since it is the main modality used in face-to-face human communication. However, pointing facilities are needed for interacting with graphical user interfaces (GUIs) efficiently; the mouse fulfills this function for direct manipulation, the prevailing paradigm for interacting with GUIs. Our work on speech and gaze human-computer interaction is based on the following observation. In some contexts of use, for instance, interaction with large displays (e.g., electronic boards/walls, reality centers, etc.) or with "Ambient Intelligence" environments [17], the mouse cannot be used, and gaze direction has the same expressive power as pointing hand gestures for designating virtual objects. In these contexts, both modalities can only specify directions, if used spontaneously as in real life; however, gaze is quicker, more precise, and less tiring than hand gestures.
Our current work on multimodality addresses the main following issue. How to design multimodal command languages that use information on spontaneous or controlled gaze movements to disambiguate oral commands, especially those including deictic phrases and ambiguous nominal references to elements in the current, virtual or real, scene?

This year, we have completed the analysis of a corpus collected earlier, with a view to gaining an insight into users' gaze strategies during oral and multimodal interaction with 3D virtual environments. This corpus includes realistic data on spontaneous and controlled eye movements concomitant with speech commands. Participants had to use a restricted vocabulary, but there was no other constraint on their spontaneous oral expression. A gender balanced group of 4 participants interacted during half an hour with various 3D applications, using speech first, then, multimodal (speech+gaze) commands. Participants could move their eyes freely during the first four tasks, while they had to look at the displayed objects they wanted to act upon during the fifth task, since they could not refer to virtual objects on the screen unambiguously using speech, due to the intentionally small size of the vocabulary. Animations of 3D objects were created using the oRis virtual reality development tool, and the user interface was simulated using an advanced implementation of the Wizard of Oz technique (i.e., the human wizard benefited from appropriate software assistance for simulating the multimodal interface). Recorded multimodal interactions have been analyzed using a specific software tool which we developed (under Linux). This tool records and "replays" interactions with any oRis application in two separate windows. One window displays the user's points of gaze superimposed on the successive displays generated by the application. The other window displays graphical representations of the speech signal and temporal evolution of the pupil diameter. It also displays the names of the graphical application objects looked at by the user (automatic labeling), as well as automatic speech recognition results in both orthographic and phonetic forms. All these data (including oral commands) are "replayed" simultaneously (fine synchronization). In addition, phonetic and orthographic transcripts of participants' spontaneous speech utterances were performed by an expert phonetician. Eye movements were also analyzed and annotated by two experts who determined (98% of agreement) which virtual objects participants had fixated successively during execution of the scenarios; fixations during speech commands and fixations preceding or following them (2 seconds before, 1 second after) were given particular attention. Results of these analyzes will be detailed and discussed in the PhD thesis that D. Gepner will defend in 2009.
The next step will be to design a real time algorithm that can, robustly and accurately: (i) interpret fixations occurring simultaneously with speech commands (loose concomitance) as designation gestures of the displayed objects on which actions mentioned in these commands are to be carried out; (ii) use this information for solving ambiguous linguistic references and deictic phrases included in oral commands (multimodal fusion). Then, this algorithm will be integrated into a software demonstrator with a view to assessing its actual efficiency (accuracy and run time speed) in realistic virtual reality environments, and testing the acceptability of the constraints forced upon users' spontaneous gaze movements and oral expression.

5.4.4. 3D hand gestures as a substitute for direct manipulation

Participants: J. Simonin, G. Pruvost, N. Carbonell.

Within the framework of the 'Modelisation, Interaction, Simulation' project supported by the CPER-Lorraine (2007-2009), the team, in cooperation with researchers (bio-physicians) of the eDAM team (UMR 7565), is designing and implementing an appropriate user interface for interacting with 3D visualizations of large molecular systems (several hundreds of thousands of items) in two situations where the use of mouse and keyboard is awkward: (i) the presentation, on a large screen, of a complex molecular system to a group of students, and (ii) the discussion of such a system with a few colleagues. The visualization software is VMD (Visual Molecular Dynamics), a free software capable of generating stereoscopic 3D views of dynamic molecular models. The input modality chosen is 3D hand gestures. For capturing finger and wrist 3D movements, we use a data glove (Immersion Cyberglove) with 18 sensors for measuring joint angles, and a magnetic sensor for determining the instantaneous position of the wrist in space (3D coordinates). Movements are sampled at a frequency of 64Hz.
This year, based on the observation and analysis of the use of VMD by eDAM researchers [4], a first 3D gesture language was designed; it enables to manipulate the visualizations generated by VMD (translations, rotations, zoom) intuitively. As this simple gestural language, which was tested by eDAM members, was intended for assessing the actual precision of motion capture sensors, it used only finger movements, the magnetic sensor

[4] Their interactions with VMD during presentations of molecular systems to undergraduate students (12 hours or so) were videotaped and analyzed.

detecting wrist position in space was ignored. The second step consisted in defining an appropriate approach for achieving reliable gesture recognition and interpretation. The approach detailed in [19] includes three main processing stages:

- computation of primitives from the analysis of digital data from the glove;
- synchronous combination of primitives to form elementary gestures (using a 50 ms window);
- asynchronous combination of elementary gestures to form valid commands.

Primitives and elementary gestures are defined as constraints on numerical values; they are specified using fuzzy logic predicates. Measures from the sensors on the glove are first preprocessed, then primitives are computed on the results of this preprocessing. A Petri net (see [28] and [31]) is used to interpret elementary gestures inferred from primitives into valid commands. The software architecture of the gesture recognizer comprises five hierarchically ordered layers of modules, one layer per type of processing: preprocessing, detection of primitives, identification of elementary gestures, recognition of valid commands, transmission of valid commands to VMD.
A generic software platform has been developed in C++ for facilitating the implementation of this approach. Gestural language designers have only to specify (in textual form) the constraints defining primitives, elementary gestures and valid commands of the language. The platform includes a debugger that displays timestamped logs of recognized primitives, elementary gestures and commands. A client-server architecture has been implemented for taking charge of exchanges between the gesture recognition system (Windows) and VMD (Linux). The initial gestural language and its usability evaluation are described in [19] which also details the adopted gesture recognition approach and the software platform.
Future work will be focused on the definition of an appropriate gestural language for manipulating VMD menus in addition to visualizations. This language will use the expressive power of both finger and hand 3D movements by interpreting the measures delivered by the magnetic sensor on the wrist. To ensure their usability, gestural commands will be designed from iconic, mimetic and deictic gestures (see the classification by Rimé [29]; semiotic gestures (as defined by Cadoz in [22]) used in everyday life will also be a source of inspiration. Then, in parallel with the ergonomic evaluation of this gestural language, issues of the design of multimodal languages appropriate for the two chosen contexts of use will be addressed. The following combinations of modalities will be experimented first: gestures and gaze; gestures and speech.

6. Other Grants and Activities

6.1. National projects

- Participation to a multi-disciplinary research project supported by the 'Fondation de France' on autism; other partners include computer scientists (at UPR 3251, LIMSI), psychologists (at UMR 7593, 'Centre Emotion'), and clinician physicians (at the 'Centre Hospitalier' in Chartres).
- Coordination of the MODIS[5] research project selected in March 2007, within MISN, a regional Research Program[6] (Contrat de plan Etat - Région Lorraine 2007-2013). MODIS aims at providing biophysicists with an appropriate multimodal (3D hand gestures, gaze and/or speech) user interface for interacting with 3D stereoscopic visualizations of molecular models and their temporal evolution, on large displays (especially reality centers or caves). Three research groups are involved in this multi-disciplinary project: eDAM (biophysics, Nancy), ETIC (ergonomics, Metz) and MErLIn.

6.2. Networks and international working groups

Participation: N. Carbonell.

- ACM Special Interest Groups: SIGACCESS (Board Officer), SIGCHI (Member).
- Scientific Committee of the CVCE Research Center, Luxembourg (Member).

[5]MODIS: Modélisation interactive de systèmes biologiques complexes en réalité virtuelle immersive.
[6]MISN: Modélisation, Information et Systèmes Numériques.

7. Dissemination

7.1. Animation of the scientific community

7.1.1. Editorial Boards of Journals

Participation: N. Carbonell.

- ACM Transactions on Accessibility. (Member of the Editorial Board).
- International Journal of Universal Access in the Information Society, Springer (Member of the Editorial Board).
- International Journal of Intelligent Information and Database Systems, Inderscience (Member of the Editorial Board).
- International Journal of Ambient Computing and Intelligence, IRMA, IGI-Global, (new journal) (Member of the Editorial Board).
- Le Travail Humain, PUF (Membre du Comité de Consultants).
- Journal d'Interaction Personne-Système, AFIHM, (new journal) (Membre du Comité Editorial).
- IEEE Transactions on Audio, Speech and Language Processing (Review).
- Interacting with Computers, Elsevier (Reviews).

7.1.2. Conference Program Committees

Participation: N. Carbonell.

- ACM 10th International Conference on Computers and Accessibility (ASSETS 2008), Halifax, Canada, October 13-15, 2008 (Steering Committee and Program Committee Member).
- ACM 10th International Conference on Multimodal Interfaces (ICMI 2008), Chania, Crete, October 20-25, 2008 (Program Committee Member: N. Carbonell; Review: O. Christmann).
- IARIA 4th Advanced International Conference on Telecommunications, Athens, Greece, June 8-13, 2008, IEEE Proceedings (Program Committee Member).
- IASTED International Conference on Assistive Technologies (AT 2008), Baltimore, Maryland, April 16-18, 2008, ACTA Press Proceedings, Canada (Program Committee Member).
- 6th International Conference on Multimedia and Network Information Systems, Wroclaw, Poland, September 18-19, 2008, IOS Press Proceedings (Program Committee Member).
- IADIS Multi Conference on Computer Science and Information Systems MCCSIS 2008) - Interfaces and Human-Computer Interaction (IHCI 2008), Amsterdam, Netherlands, July 22-27, 2008 (Program Committee Member).
- 4th Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT), Cambridge, UK, April 13-16, 2008, Springer- Verlag Proceedings (Program Committee Member).
- KES 2nd International Symposium on Agent and Multi-agent Systems: Technologies and Applications (KES-AMSTA 2008), Incheon, Korea, March 27-28, 2008, Springer LNCS Proceedings (Program Committee Member).
- 4th Usability Symposium of the WG HCI&UE of the Austrian Computer Society (USAB 2008), Graz, Austria, November 20-21, 2008, Springer LNCS Proceedings (Program Committee Member).
- Workshop on Human-computer Interaction for Medicine and Health Care (HCI4MED), at 22th British HCI Group Annual Conference, September 1-5, 2008, Liverpool, UK, publication of selected papers in a special issue of the International Journal of Human-Computer Studies (IJHCS) (Program Committee Member).

- 20ième Conférence Francophone sur l'Interaction Homme-Machine, Metz, France, 2-5 Septembre, 2008, Actes (abstracts): ACM Digital Library. (Program Committee Co-chair: N. Carbonell; Reviews: O. Christmann, J. Simonin).
- 11ième Conférence ERGO-IA, Ergonomie et Informatique avancée, Biarritz, France, 15-17 octobre 2008 (Program Committee Member).
- 3ième Workshop sur les Agents Conversationnels Animés (WACA 2008), Paris, 28 novembre 2008 (Program Committee Member)

7.1.3. Others

Expert: N. Carbonell

- Evaluation of a European Integrated Project: First year review of the FP6 IP COMPANIONS (January 29-30, 2008, Sheffield, UK).
- Evaluation of a network of excellence proposal for the Austrian National Research Network (NFN): review of the proposal'Diversability - Technology for All' (Vienna, Austria, September 24-25, 2008.
- Evaluation of the application of Dr. Dimitris Grammenos for an Assistant Researcher position (i.e., Researcher C) at The Institute of Computer Science (ICS) of the Foundation for Research and Technology - Hellas (FORTH).

7.1.4. Ph.Ds and Habilitations examining boards

Participation à des Jurys: N. Carbonell

- Thomas Gaudy. Etude et développement de jeux vidéo sonores accessibles aux personnes aveugles. Thèse du Conservatoire National des Arts et Metiers (CNAM), spécialité informatique, Paris, 3 juillet 2008. (Rapporteur).
- Stephan Raidt. Gaze and face-to-face communication between a human speaker and an animated conversational agent - Mutual attention and multimodal deixis. Thèse de l'Institut National Polytechnique de Grenoble, spécialité 'Signal, Image, Parole, Télécoms', Grenoble, 2 avril 2008. (Examinateur).
- Blaise Potard. Inversion acoustique-articulatoire avec contraintes. Thèse de l'Université Henri Poincaré, spécialité informatique, Nancy, 23 octobre 2008. (Examinateur).

7.2. Teaching

- Master Informatique, Universités de Nancy, M1, U.E. d'ossature 'Modèles de perception et raisonnement': N. Carbonell (24h).
- Master Informatique, Universités de Nancy, M1, U.E. de différenciation 'Conception et évaluation d'interfaces utilisateur': N. Carbonell (30h).
- Master Informatique, Universités de Nancy, M2, spécialité à finalité recherche 'Perception, Raisonnement, Interactions Multimodales', U.E. de différenciation 'Modélisation de l'utilisateur et flexibilité des interfaces': N. Carbonell (15h).
- Master Information et Communication, Université Nancy 2, spécialité à finalité professionnelle 'Information Scientifique et Technique et Intelligence Economique', UE d'ossature 'Algorithmique et programmation': N. Carbonell (20h).
- Master Informatique, Universités de Nancy, Responsabilité de la spécialité recherche 'Perception, Raisonnement, Interactions Multimodales': N. Carbonell.
- Master Informatique, Universités de Nancy, Membre de l'Equipe de formation du Master et membre permanent du jury de soutenance des stages (M2) pour les 4 spécialités à finalité recherche : N. Carbonell.

- Master Information et Communication, Université Nancy 2, Co-responsabilité de la spécialité à finalité professionnelle 'Information Scientifique et Technique et Intelligence Economique': N. Carbonell.
- 2nd and 3rd year (Master), Filière 'Ingénierie des Systèmes Automatisés', Ecole Nationale Supérieure d'Electricité et de Mécanique, Nancy : Real time, Embedded systems: O. Christmann (190h), D. Gepner (90h).

7.3. Participation to conferences, workshop, invited talks

- First year review of the European FP6 IP COMPANIONS, January 29-30, 2008, Sheffield, UK. (N. Carbonell).
- 10ièmes Rencontres Internationales de la Réalité Virtuelle (Laval Virtual 2008), Laval, France, 9-13 avril 2008. (O. Christmann, D. Gepner, J. Simonin).
- Meeting of the Scientific Committee of the Centre Virtuel de la Connaissance sur l'Europe (CVCE), June 27, 2008. (N. Carbonell).
- 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2008), Hannover, Germany, July 29 - August 1, 2008. (J. Simonin).
- 20ième Conférence Francophone sur l'Interaction Homme-Machine, Metz, France, 2-5 Septembre, 2008. (N. Carbonell, O. Christmann).
- 10th International Conference on Visual Information Systems (Visual 2008), Salerno, Italy, September 11-12, 2008. (N. Carbonell, O. Christmann).
- Evaluation of a network of excellence proposal ('Diversability - Technology for All') for the Austrian National Research Network (NFN): review meeting, Vienna, Austria, September 24-25, 2008. (N. Carbonell).
- ACM 10th International Conference on Computers and Accessibility (ASSETS 2008), Halifax, Canada, October 13-15, 2008. (N. Carbonell)
- ACM 10th International Conference on Multimodal Interfaces (ICMI 2008), Chania, Crete, October 20-25, 2008. (N. Carbonell, J. Simonin).
  Invited talk at a panel:'Multi-modal: Because we can or because we should?' (N. Carbonell).
- 3ième Workshop sur les Agents Conversationnels Animés (WACA 2008), Paris, 28 novembre 2008. (N. Carbonell).

8. Bibliography

Major publications by the team in recent years

[1] A. CAPOBIANCO. *Questioning the effectiveness of contextual online help: some alternative propositions*, in "Proceedings of 9th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT'03)", M. RAUTERBERG, M. MENOZZI, J. WESSON (editors), Amsterdam: IOS Press, Zürich, Switzerland, September, 1-5, 2003, p. 65-72.

[2] A. CAPOBIANCO, N. CARBONELL. *Online help for the general public: specific design issues and recommendations*, in "International Journal of Universal Access in the Information Society (UAIS), Springer", vol. 2(3): 265-279, 2003.

[3] N. CARBONELL. *Recommendations for the design of usable multimodal command languages*, in "International Journal of Universal Access in the Information Society (UAIS), Springer, Special issue 'Multimodality, a step towards universal access'", vol. 2/2: 143-159, 2003.

[4] N. CARBONELL, S. KIEFFER. *Do oral messages help visual search?*, in "Advances in Natural Multimodal Dialogue Systems", J. VAN KUPPEVELT, L. DYBKJAER, N. BERNSEN (editors), vol. 30, Dordrecht (NL): Springer, "Text, Speech and Language Technology" Series, 2005, p. 131-157.

[5] O. CHRISTMANN, N. CARBONELL. *Browsing through 3D representations of unstructured picture collections: an empirical study*, in "Proceedings ACM Working Conference on Advanced Visual Interfaces", A. CELENTANO, P. MUSSIO (editors), New York: ACM Press, Venezia, Italy, May 23-26, 2006, p. 445–448.

[6] D. GEPNER, J. SIMONIN, N. CARBONELL. *Gaze as a supplementary modality for interacting with ambient intelligence environments*, in "Proceedings of 12th International Conference on Human-Computer Interaction (HCI International)", C. STEPHANIDIS (editor), LNCS-LNAI, vol. 4555, 'Universal Access to Ambient Interaction', Springer-Verlag, Beijing, China, July 22-27, 2007, p. 848-857.

[7] S. KIEFFER, N. CARBONELL. *How really effective are multimodal hints in enhancing visual target spotting? Some evidence from a usability study.*, in "International Journal on Multimodal User Interfaces (JMUI)", vol. 1(1): 1-5, 2007.

[8] J. SIMONIN, N. CARBONELL. *Interfaces Adaptives - Adaptation dynamique à l'utilisateur courant*, in "Interfaces numériques", I. SALEH, D. REGOTTAZ (editors), Paris: Hermès-Lavoisier, 2007, p. 37-54.

[9] J. SIMONIN, S. KIEFFER, N. CARBONELL. *Effects of display layout on gaze activity during visual search*, in "Proceedings of 10th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT'05), LNCS Series num: 3585", M.-F. COSTABILE, F. PATERNO (editors), LNCS, vol. 4555, Springer-Verlag, Rome, Italy, September 12-16, 2005, p. 1054-1058.

[10] J. SIMONIN. *Aide en ligne adaptative et assistants conversationnels animés : mise en oeuvre et évaluation ergonomique*, Ph. D. Thesis, Université Henri Poincaré, Nancy, October 2007.

[11] F. VALDENAIRE. *Modulation de la précision de l'affichage en fonction de la direction du regard*, 23 p., Mémoire de stage de recherche, Master Informatique des Universités de Nancy, LORIA, Septembre 2007.

Year Publications

Doctoral Dissertations and Habilitation Theses

[12] O. CHRISTMANN. *Navigation dans de grands ensembles structurés de documents visuels*, Ph. D. Thesis, Université Henri Poincaré, Nancy, October 2008.

International Peer-Reviewed Conference/Proceedings

[13] O. CHRISTMANN, N. CARBONELL. *Navigation and Search in 3D Visualizations of Large Unstructured Photo Collections: an Empirical Study*, in "Proceedings of 10th International Conference on Visual Information Systems (Visual 2008)", M. SEBILLO, G. VITIELLO, G. SCHAEFER (editors), LNCS, vol. 5188, Springer-Verlag, Salerno, Italy, September 11-12 2008, p. 31–43.

[14] J. SIMONIN, N. CARBONELL, D. PELÉ. *Effectiveness and usability of an online help agent embodied as a talking head*, in "Proceedings of 10th International Conference on Multimodal Interfaces (ICMI 2008)", V. DIGALAKIS, A. POTAMIANOS, M. TURK, R. PIERACCINI, Y. IVANOV (editors), ACM DL, Chania, Crete, Greece, October 20-22 2008, p. 17–20.

[15] J. SIMONIN, N. CARBONELL. *Proactive Versus Multimodal Online Help: An Empirical Study*, in "Proceedings of 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2008)", W. NEJDL, J. KAY, P. PU, E. HERDER (editors), LNCS, vol. 5149, Springer-Verlag, Hannover, Germany, July 29 - August 1, 2008, p. 183-192.

National Peer-Reviewed Conference/Proceedings

[16] J. SIMONIN, N. CARBONELL, D. PELÉ. *Apports d'une tête parlante à l'utilisation d'aides en ligne grand public ?*, in "Actes du 3ième Workshop sur les Agents Conversationnels Animés (WACA 2008)", to appear, Paris, 28 novembre 2008, 7.

Scientific Books (or Scientific Book chapters)

[17] N. CARBONELL. *Chapter 51 - Contributions of "ambient" multimodality to Universal Access*, in "The Universal Access Handbook", C. STEPHANIDIS (editor), to appear, Taylor & Francis Group, 2008, 35.

Other Publications

[18] N. CARBONELL. *Multi-modal: Because we can or because we should?*, (Invited talk, Panel, at 10th International Conference on Multimodal Interfaces (ICMI 2008), Chania, Crete, Greece, October 20-22, 2008).

[19] G. PRUVOST. *Modélisation du geste pour la manipulation de visualisations 3D en environnement immersif*, Juin 2008.

[20] J. SIMONIN. *Software Platform for Assisted Simulation of Novel User Interfaces, Replay and Annotation of Rich Multimodal Interaction Logs*, (Demonstration at 10th International Conference on Multimodal Interfaces (ICMI 2008), Chania, Crete, Greece, October 20-22, 2008).

References in notes

[21] B. B. BEDERSON. *PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps*, in "CHI Letters", vol. 3(2): 71-80, 2001.

[22] C. CADOZ. *Le geste canal de communication homme-machine, la communication instrumentale*, in "Technique et Science Informatiques (TSI)", vol. 13(1): 31-61, 1994.

[23] M. DORR, T. MARTINETZ, K. GEGENFURTNER, E. BARTH. *Guidance of Eye Movements on a Gaze-Contingent Display*, in "Proceedings of 5th Workshop on Dynamic Perception", U. ILLG, H. BÜLTHOFF, H. MALLOT (editors), IOS Press, Tübingen, Germany, November 18-19, 2004, p. 89-94.

[24] A. T. DUCHOWSKI. *Eye Tracking Methodology: Theory and Practice*, 328 p., Springer, 2007.

[25] T. GRINDINGER. *Eye Movement Analysis and Prediction with the Kalman Filter*, 19 p., Master thesis (Advisor: A. T. Duchowski), Clemson University, SC, 2006.

[26] R. JACOB. *Eye tracking in advanced interface design*, in "Virtual Environments and Advanced Interface Design", W. BARFIELD, T. FURNESS (editors), New York: Oxford University Press, 1995, p. 258-288.

[27] A. JAMESON. *Adaptive Interfaces and Agents*, in "The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications", J. JACKO, S. SEARS (editors), Mahwah, NJ: Lawrence Erlbaum, 2003, p. 305-330.

[28] S. KURKOVSKY, R. LOGANANTHARAJ. *Extension of Petri Nets for Representing and Reasoning with Tasks with Imprecise Durations*, in "Applied Intelligence", vol. 23(2): 97-108, 2005.

[29] B. RIMÉ. *A trip to the overlooked world of hand gestures*, in "Contemporary Psychology", vol. 39(1): 55-56, 1994.

[30] G. ROBERTSON, M. CZERWINSKI, K. LARSON, D. C. ROBBINS, D. THIEL, M. VAN DANTZICH. *Data Mountain: Using Spatial Memory for Document Management*, in "Proceedings of 11th ACM Annual Symposium on User Interface Software and Technology (UIST'98)", New York: ACM Press, San Francisco, CA, November 1-4, 1998, p. 153-162.

[31] M. ROMBAUT, I. JARKASS, T. DENOEUX. *State Recognition in Discrete Dynamical Systems Using Petri Nets and Evidence Theory*, in "Proceedings of 5th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, (ECSQARU)", LNCS, vol. 1638, Springer-Verlag, London, UK, July 5-9 1999, p. 352–361.