Natural User Interface (NUI): a case study of a video based interaction technique for a computer game

M. Rauterberg

Institute for Hygiene and Applied Physiology (IHA)
Swiss Federal Institute of Technology (ETH)
Clausiusstrasse 25, CH-8092 Zurich, Switzerland
http://www.ifap.bepr.ethz.ch/~rauterberg

To compare the advantages and disadvantages of a *Natural User Interface* [1] a field study was carried out. During five days of the largest computer fair in Switzerland four different computer stations with (1) a command language, (2) a mouse, (3) a touch screen, and (4) a Digital Playing Desk (DPD) interface was presented for public use. In this version of the DPD the user has to play a board game by moving a real chip on a virtual playing field against a virtual player. The task was to win the computer game "Go-bang". The reactions of the virtual player were simulated by "emoticons" as colored comic strip pictures with a corresponding sound pattern. We investigated the effects of these four different interaction techniques with an inquiry with a questionnaire. Results of the inquiry: 304 visitors rated the usability of all four stations on a bipolar scale. The touch screen station was rated as the easiest to use interaction technique, followed by the mouse and DPD interface; the "tail-light" was the command language interface. One very important result was a significant correlation between "age" and "DPD usability". This correlation means that older people prefer significantly more a graspable user interface in form of the DPD than younger people.

1. INTRODUCTION

There are two main contrary directions for new interface technology: (1) [immersive] virtual reality, and (2) augmented reality or ubiquitous computing. Enthusiasts of the virtual reality approach (VR) believe "that all constrains of the real world can be overcome in VR, and physical tools can be made obsolete by more flexible, virtual alternatives" ([2] p. 87-88). We restrict the notion of "virtual reality system"--in the context of this paper--to systems with head mounted display and data gloves or suits. In these VR applications the user has to leave his natural physical and social environment and to *immerse* in the simulated world. The following two unsolved problems are important: (1) how to simulate tactile and haptic feedback, and (2) how to overcome the social isolation for collaborative tasks. The effect, that the social nearness between real persons is of tremendous importance for collaboration, was investigated and shown in [3]. As we stated in [4], that the effects of tactile and haptic senses are important, we are looking for a realization of a user interface where the user can control the human-computer interaction by his hands dealing with real and virtual objects in the same interface space. The DigitalDesk of Wellner [2] was one of such systems.

Inspired by the ideas of Wellner [2], we were interested in a way to test empirically the advantages or disadvantages of the DigitalDesk in comparison with established interaction techniques. The DigitalDesk has the following three important features: (1) it projects electronic images (virtual objects) down onto the desk and onto real objects, (2) it responds to interaction with real objects (e.g., pens or bare fingers: hence *Digital*Desk), and (3) it can interpret the scene on an appropriate semantic level (e.g., read paper documents placed on the desk; cf. [2]).

2. SYSTEM DESCRIPTION

To run a laboratory investigation or a field study we need a fast, reliable and robust implementation of the whole system. First, we decided to minimize the task complexity and to restrict the user's action space to cognitive planning processes. For public use a simple computer game seems to be best. We implemented a version of the computer game "Gobang". The user has to play the game by moving a *real* chip on the *virtual* playing field (see Fig. 1). To compare this interface type with the most established dialog techniques we implemented the same game algorithm on three other stations with (1) a command language, (2) a mouse, and (3) a touch screen interface.

Command interface (CI): This station run on a 386er PC with a color screen (17") in an upright position. The user has to enter the co-ordinates of the desired place of a playing field with 12 by 12 positions (e.g., A1, L12). To start a new or to cancel a game she or he has to enter the command NEW. The internal state of the algorithm is presented as text in a special output field (e.g., "Make the next move").

Mouse interface (MI): This interface run on a 386er PC with a color screen (17") in an upright position. To move the user has to click with the mouse on the desired place. To start a new or to cancel a running game she or he has to click on the button NEW. The internal state of the algorithm is presented as text in a special popup window.

Touch screen interface (TI): This station run on 386er PC with a color touch screen (21") in an inclined position of 30 degrees. To make a move the user has to touch with a finger the desired place. To start a new or to cancel a running game she or he has to touch the button NEW. The virtual player was shown on a second colored screen (17") served by a second 386er PC in a client-server architecture.

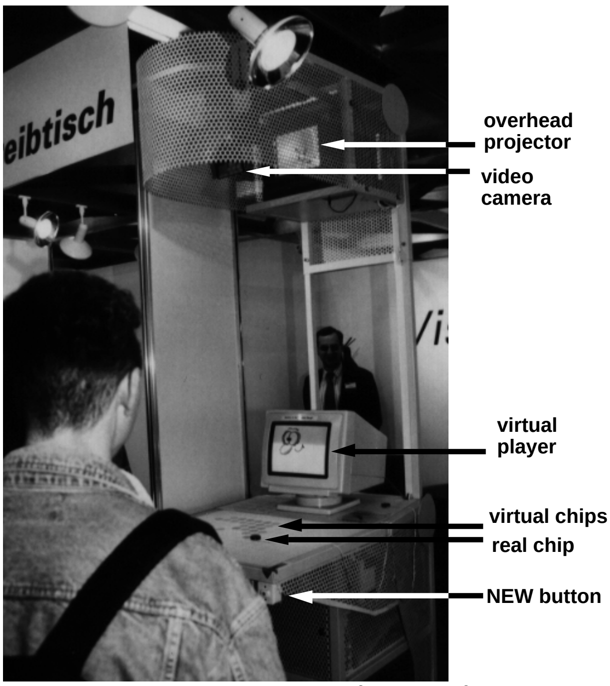

Fig. 1: The Digital Playing Desk--the front view for the users.

Fig. 2: The two emoticons "reasoning" and "yawning".

Digital Playing Desk (DPD): This station was completely realized in C^{++} on standard hardware components: (1) a Pentium PC, (2) an overhead projector of high luminous intensity and the projection panel GehaVision™, (3) a high resolution video camera, and (4) the video board MovieMachinePro™. For the virtual player a second 386er PC was connected in a client-server architecture. A user has to make a move by putting a *real* chip on the desired place of the *virtual* playing field (see Fig. 1). The computer's output is the projection of a *virtual* chip on the desk. If a user wanted to cancel or restart a game, then he or she had to press the real NEW button in the front of the station.

The Virtual Player: The CI and MI have their output screens in an upright position. This upright position makes it impossible to give an opponent--in the metaphor of a game--an individual representation. The classical solution is, to give feedback about the machine's internal states as text or graphics in defined areas on the screen. This solution leads always to a partition of the screen into a working area (e.g., the playing field) and the feedback area. This superimposing of qualitatively different feedback's in the same output space can be overcome, if we use an additional output device. This can be done, if we separate the working area (the playing field) from the feedback area and if we change the upright position of the working area to a horizontal position. Therefore we composed the system of a flat table to project the playing field onto and of a second screen to present the virtual player.

All--from the user's point of view--important internal states of the game algorithm were presented by the virtual player with six different facial expressions and a corresponding sound (see Fig. 2; partially animated comic strip pictures: "Reasoning" = a face with a balloon of animated turning wheels and a machine-like sound, "Waiting for the next move" = a yawning face with a corresponding sound, "Initial state" = a face with blinking eyes, "Incorrect move" = an angry facial expression with an indignant cry, "Be the winner" = a happy facial expression with an arrogant laughter, "Be the loser" = a shrinking face with a disappointed cry). We call these six comic strip pictures "emoticons" (acronym for "emotional icons"). All emoticons of the virtual player are shown on a second color screen (17").

3. VALIDATION OF THE DIGITAL PLAYING DESK

To present our four dialog techniques to a broad population of heterogeneous users, four special stations were constructed. All stations were presented in a central exhibition area during five days at the largest computer fair of Switzerland in September 1995. The official number of visitors was approximately 70'000. Most of these people passed the exhibition area and many of them came into close contact with one of the stations (e.g., playing at one of the stations or observing other people playing). With a questionnaire we got some personal information's of several users of at least one of our four stations. The stand personal was instructed to request users to answer the questionnaire. It was especially necessary to ask women, because they behaved very reserved.

Subjects: The questionnaire was answered by 304 visitors (61 females, 243 males, 5 anonymous data). The average age of the women was 31 ± 13 years, and of the men 30 ± 14 years (T-Test: $p \leq .724$). As a *control variable* the "computer experience" was measured in millimetre on the corresponding bipolar rating scale ["no experience": 0 … 90 mm: "expert"]. *There* was a significant gender difference in computer experience; the computer experience of men was higher than the experience of women (men: 58 ± 22 mm; women: 48 ± 23 mm; T-Test: $p \leq .002$).

Dependent Measures: The questionnaire consisted of two parts: (a) personal data (age in years, gender, computer experience in form of a bipolar rating scale, and (b) usability ratings. For each of the four interfaces the following five aspects were asked with a multiple choice question: "did you play", "did you loose", "did you win", "did you play draw", "did you cancel". The *dependent measure* per interface type is the number of millimetres of the user's marking on the bipolar rating scale ["very easy to use": 0 … 70 mm: "very difficult to use"]. We also differentiated between "real station contact" and "only observer status".

Results: All persons, who answered that they had no real contact to one of the stations, had a significant higher amount of computer experience ("no contact": 70 ± 17 mm, N = 10; "with contact": 55 ± 23 mm, N = 276; T-Test: $p \leq .04$). We can not find a significant difference between the "contact" and "no contact" group in the usability ratings.

The usability rating was best for the touch screen interface (median = 3; mean rank = 1.9), followed by the mouse interface (median = 6; mean rank = 2.4) and the Digital Playing Desk (median = 9; mean rank = 2.5). The "tail-light" was the command language interface (median = 17; mean rank = 3.3). These differences are significant (Friedman Test, df = 3, Chi^2 corrected for ties = 89.1, $p \leq .0001$). One very important result was a significant correlation between "age" and "Digital Playing Desk's usability". This correlation means that older people prefer significantly more a graspable user interface in form of the Digital Playing Desk than younger people. ($R = -0.202$, $N = 179$, $p \leq .006$). All other correlations are not significant.

4. DISCUSSION AND CONCLUSION

To carry out an inquiry with a questionnaire only for scientific purposes in the context of a commercial fair was more difficulty than we expected. The only motivation for the user to fill out such a questionnaire--in contrast to all the lotteries of the commercial issuers around us (with their very attractive prizes)--was to bring in his or her personal opinion into an scientific research process. This argumentation was the most convincing reason to participate. Overall we got more filled out questionnaires than we expected, but much less than observed visitors at one of the four stations. To increase the number of answered questionnaires the stand personal has to be active: to go to the user and to ask for participation.

We could find two main results: (1) the touch screen interface was estimated as the easiest to use, and (2) the significant correlation between age and the usability ratings for the Digital Playing Desk. If we assume, that the average age of the populations in all high industrialised countries will increase in the next two or three decades, then this result will be of tremendous importance for the development of modern computer technology for elderly people!

The general advantage and disadvantage of immersive VR are the necessity to put the user into a complete modelled virtual world. This concept of immersing the user in the computer's world ignores the on-going process of interacting with the real world. In the same interface space the mixing of real and virtual objects is not possible. But, humans are--most of their time--part of a real world and interact with real objects and other real humans.

Augmented Reality (AR) recognises that people are used to the real world and that the real world cannot be reproduced completely and accurately enough on a computer. AR builds on the real world by augmenting it with computational capabilities. AR is the general design strategy behind a "Natural User Interface" (NUI) [1].

A system with a NUI supports the mix of real and virtual objects in the same interaction space. As input it recognises *and* understands physical objects and humans acting in a natural way (e.g., object handling, hand writing, etc.). Its output is based on pattern projection such as video projection, holography, speech synthesis or 3D audio patterns. A necessary condition in our definition of a NUI is that it allows inter-referential I/O, i.e. that the same modality is used for input *and* output (see [4]). For example, a projected item can be referred directly by the user for his or her nonverbal input behavior.

REFERENCES

[1] Rauterberg M & Steiger P, Pattern recognition as a key technology for the next generation of user interfaces. In Proc. of SMC'96 (Vol. 4, IEEE Catalog Number: 96CH35929, pp. 2805-2810). Piscataway: IEEE (1996).

[2] Wellner P, Interacting with paper on the DigitalDesk. In *Com. of the ACM*. 36(7), pp. 86-96 (1993).

[3] Rauterberg M, Dätwyler M & Sperisen M, The shared social space as a basic factor for the design of group-ware. In: K. Brunnstein & P. Sint (Eds.) Intellectual Property Rights and New Technologies. (Schriftenreihe der Österreichischen Computer Gesellschaft Vol. 82, pp. 176-181), Oldenbourg (1995).

[4] Rauterberg M & Szabo K, A Design Concept for N-dimensional User Interfaces. In Proc. of 4th Intern. Conf. INTERFACE to Real & Virtual Worlds (1995) pp. 467-477.

Advances in Human Factors/Ergonomics, 21B

Design of Computing Systems:

Social and Ergonomic Considerations

Proceedings of the Seventh International Conference on Human-Computer Interaction, (HCI International '97), San Francisco, Califormia, USA
August 24-29, 1997
Volume 2

Edited by
Michael Smith
University of Wisconsin, Madison, WI 53706, USA
Gavriel Salvendy
Purdue University, West Lafayette, IN 47907, USA
Richard J. Koubek
Purdue University, West Lafayette, IN 47907, USA

ELSEVIER SCIENCE **1997**
Amsterdam – Lausanne – New York – Oxford – Shannon – Tokyo