Natural User Interfaces (NUI): a case study of a video based interaction technique for CAD systems

M. Rauterberg[a], M. Bichsel[b], H. Krueger[a] & M. Meier[b]

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

It is time to go beyond the established approaches in human-computer interaction. We discuss the two known approaches to overcome the obstacles and limitations of traditional interfaces: [immersive] Virtual Reality (VR), and Augmented Reality (AR). Following the fundamental constraints of the natural way of interacting with real world objects we derive a set of recommendations for the next generation of user interfaces: the *Natural User Interface* (NUI). The concept of NUI is discussed in form of a general framework and in form of an implemented prototype: a video-based interaction technique for a planning tool for construction and design tasks. Results of a first empirical evaluation study are reported.

1. INTRODUCTION

The introduction of computers into the workplace has had a tremendous impact on the field of human-computer interaction. Mouse driven graphic displays are everywhere: the desktop workstations define the frontier between the computer world and the real world. We spend a lot of time and energy to transfer information between those two worlds. This could be reduced by better integrating the virtual world of the computer with the real world of the user.

In all the traditional interaction styles (command, menu, direct manipulation) the user can not mix real world objects with virtual objects in the *same* interface space. They also do not take into considerations the enormous potential of human hands to interact with real and virtual world objects. This aspect was one of the basic ideas to develop data gloves and data suits for interactions in an immersive virtual reality system (VR). The other basic idea, to realise a VR system, was the 3D output capabilities in the usage of head mounted displays. However, in (immersive) VR systems several serious problems are inherently present:
• The lack of tactile and of touch information and consequently the mismatch with the proprioceptive feedback. Special techniques are proposed to overcome this problem.
• The lack of information for depth perception, since visual displays generate 2D data. Many concepts are generating possibilities to reconstruct 3D pictures from these 2D data [1].
• There is always a time delay in the user-computer control loop, which causes severe problems with reference to the perceptual stability of the vestibular apparatus in the ear.
• The influence of continuous communication on social interaction, based on a shared social space, is of paramount importance for computer supported cooperative work.

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 most of their time, humans are a part of the real world and interact with real objects and other real humans.

The most promising approach to overcome the obstacles of immersive VR is *Augmented Reality* (AR) [2]. The expected success of this approach lies in its ability to build on fundamental human skills: namely, to interact with real world subjects and objects! To empower the human to computer interaction, the user must be able to behave in a *natural way* bringing into action all of their body parts (e.g., hands, arms, face, head, voice, etc. [3]). To interpret all of these expressions we need very powerful and intelligent pattern recognition techniques. In this paper we present an alternative interaction style, called *Natural User Interface*.

2. A FRAMEWORK FOR A NATURAL USER INTERFACE (NUI)

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 "Natural User Interfaces". A system with a NUI supports the mix of real and virtual objects. As input it recognises (visually, acoustically or with other sensors) and understands physical objects and humans acting in a natural way (e.g., speech input, 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. For example, a projected item can be referred directly by the user with their nonverbal input behavior. Figure 1 provides an overview how a system with a NUI could look.
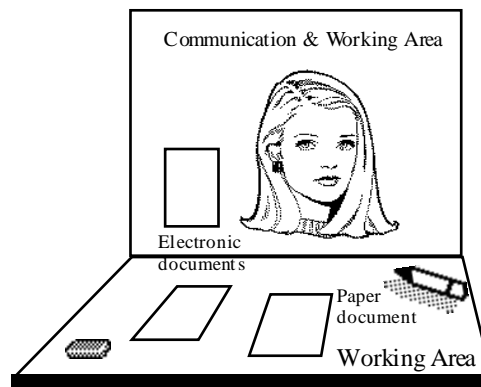

Figure 1: Architecture of a Natural User Interface (see also [4] and [5])

A set-up of several parallel input channels via video cameras allows the showing of multiple views to remote communication partners, such as a (3D) face view [4] and a view of shared work objects [2]. Multimedia output is provided through the vertical communication area display, the projection device from the top down to the working area and through [four] loudspeakers, producing a spatial impression on the user. Free space in the communication area can be used for (content) work, e.g. electronic documents (see Figure 1). Of course, traditional input and output devices still can be used in addition. As introduced by Tognazzini [5], NUIs are multimodal and therefore allow users to choose for every action the appropriate and individually preferred interaction style.

Since human beings manipulate objects in the physical world most often and most naturally with hands, there is a desire to apply these skills to human-computer interaction. NUIs allow the user to interact with real and virtual objects on the same working area in a *literally* direct manipulative way! The working area is primarily horizontal, so that user can put real objects on the surface. Users get the feedback of the state of manipulated objects exactly at the same place where they manipulate these objects: perception space and action space coincide! This very powerful design criterion was discovered by Rauterberg and empirically validated [6].

3. THE PROTOTYPE "BUILD-IT"

In a first step we designed a system that is based primarily on the concept of NUIs with the communication aspects of a computer supported cooperative work environment: the design table (see Figure 2). A multidisziplinary team can sit around the "augmented" design table to solve planning tasks in a cooperative manner. We choose the task context of planning activities for plant design. A prototype system, called "BUILD-IT", has been realized and an application, supporting engineers in designing assembly lines and building plants, was implemented. BUILD-IT enables the users to act with a mixture of virtual and real world objects in the same interaction space. The vertical area for communication and work (in the background of Figure 2) is used for a 3D-perspective view into the plant.
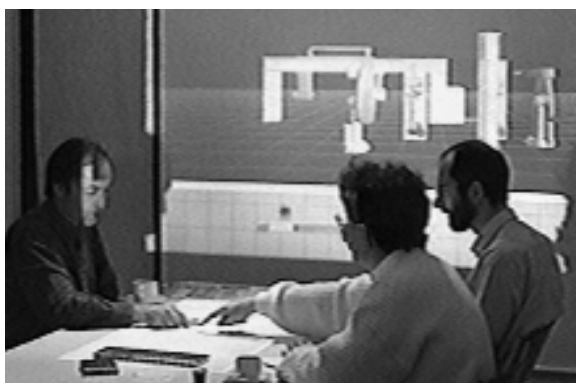
Figure 2: BUILD-IT--the augmented design table.
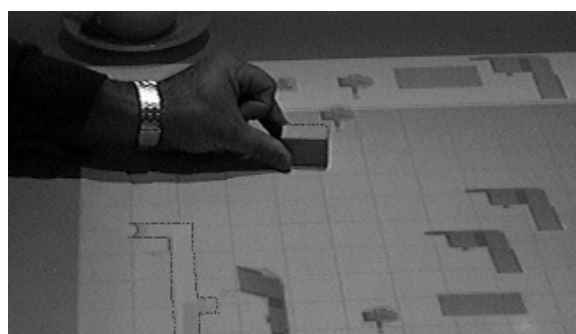
Figure 3: The top view with the planning area below (grey), the menu area above (white), and a human hand moving the interaction handler (the brick).

The system BUILD-IT has seven components:
1) A table with a white surface is used as a horizontal projection and interaction area.
2) A white projection screen provides a vertical projection area.
3) A high resolution LCD projector projects the top view vertically down to the design table.
4) A high resolution LCD projector projects a 3D-perspective view horizontally onto the projection screen in the background.
5) A high resolution CCD camera looks vertically down to the activities on the design table.
6) A small brick is the physical interaction 'device' (the universal interaction handler) that gives the user the feeling of haptic feedback of the virtual objects moved around.
7) A low-cost Silicon Graphics Indy O2 provides the computing power.
The software consists of two independent processes communicating with each other through socket connection.
A) The real time analysis process of the video reads images from the camera, extracts contours of moving objects, analyses these contours [7] [8], and determines the position and orientation of the universal interaction handler (the brick).
B) An application is built using the multi-media framework MET++ [9]. This application interprets the user's action based on the position and orientation of the interaction handler, modifies a virtual scene according to the user action.

The application is designed for supporting providers of assembly lines and plants in the early design processes. It can read and render arbitrary CAD models of machines in VRML format. The input of a 3D model of virtual objects is realized by connecting BUILD-IT with the CAD-System CATIA. This connection guarantees the import of original CAD-models into BUILD-IT. BUILD-IT currently supports the following user (inter-)actions (see Figure 3):

- Selection of a virtual object (e.g., a specific machine) in the 'virtual machine store' by placing the interaction handler onto the projected image of the machine in the menu area.
- Positioning a machine in the virtual plant by moving the interaction handler to the corresponding position in the projected plant layout on the table. Positioning includes machine orientation that is coupled to the orientation of the interaction handler.
- Fixing the machine by covering the surface of the interaction handler with the hand and removing the brick from that particular position.
- Deleting the machine by moving it back into the menu area (the virtual machine store).
- Moving a virtual camera through the plant, rendering the 3D-perspective view, and displaying it on the vertical projection screen (see Figure 2).

The system has been empirically tested with several managers and engineers from companies that produce assembly lines and plants. These tests showed that the system is intuitive to use, easy to learn, and that people could enjoy using it. Most people were able to assemble virtual plants after a few minutes of introduction to the system.

Some specific user comments are: "The concept phase is especially important in plant design since the customer must be involved in a direct manner. Often, partners using different languages sit at the same table. This novel interaction technique will be a means for completing this phase efficiently and almost perfectly". "Improvement of the interface to the customer in the offering phase as well as during the project, especially in simultaneous engineering projects". "A usage of the novel interaction technique will lead to a simplification, acceleration, and reduction of the iterative steps in the starting and concept phase of a plant construction project".

4. CONCLUSION

One of the most interesting advantages of a NUI-like interface is the possibility to mix real world objects with virtual objects in the same interaction space [4]. One of our next steps is the implementation of two or three interaction handlers to allow simultaneous interactions among several users sitting around the design table. By using this new interaction technique it will be possible to discuss and manage difficult 3D objects with any customer who might not be expert in CAD-systems. Technical descriptions and products can be presented in an easy way, and changes, affected by changing requirements, can be realized and presented in a very short time. The virtual camera allows a walk-through through the designed plant. Such kind of walk-through gives very good feedback of all parts in a complex system.

REFERENCES

[1] 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.
[2] Wellner P, Mackay W, Gold R, Computer-Augmented Environments: Back to the Real World. Communications of the ACM, 36(7) (1993) pp. 24-26.
[3] Fitzmaurice G, Ishii H, Buxton W, Bricks: Laying the Foundations for Graspable User Interfaces. In Proc. of the CHI ´95 (1995) pp. 442-449.
[4] Okada K, Ichikawa Y, Jeong G, Tanaka S & Matsushita Y, Design and evaluation of MAJIC videoconferencing system. In Proc. of IFIP INTERACT'95 (1995) pp. 289-294.
[5] Tognazzini B, Tog on Software Design. Addison-Wesley, Reading MA (1996).
[6] Rauterberg M, Über die Quantifizierung software-ergonomischer Richtlinien. PhD Thesis, University of Zurich (1995).
[7] Bichsel M, Illumination Invariant Segmentation of Simply Connected Moving Objects. 5th British Machine Vision Conference, University of York, UK (1994) pp. 459-468.
[8] Bichsel M, Segmenting Simply Connected Moving Objects in a Static Scene. Transactions on Pattern Recognition and Machine Intelligence (PAMI), 16(11) (1994) pp. 1138-1142.
[9] Ackermann P, Developing Object-Oriented Multimedia Software Based on the MET++ Application Framework. Heidelberg: dpunkt Verlag für digitale Technologie (1996).

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