paper

perceptual user interface for human-computer interaction

Weidong Geng, Vladimir Elistratov, Marina Kolesnik, Thomas Kulessa, Wolfgang Strauss
Fraunhofer Institute for Media Communication IMK, MARS-Exploratory Media Lab
D-53754, St. Augustin, Germany
Email: {geng,elistratov,kolesnik,thomas.kulessa,strauss}@imk.fhg.de

Abstract:

In this paper we present our effort towards perceptual user interface for main interaction tasks, such as navigation/travel, selection/picking and personal data access, in e-commerce environment. A set of intuitive navigation devices, including Treadmill, Virtual Balance and Cyberwheel, are described and evaluated for web/public space. Vision-based pointing is explored for free-hand selection/picking, and wireless personal access system is developed for unobtrusive transmission of personal information from hand-held devices. Furthermore, we implement an integrated interaction platform, which could couple these devices together in complex scenarios such as portals for shopping.

Keywords: Perceptual User Interface, 3D Interaction, Human-computer Interface

Project URL:
http://imk.gmd.de/mars

1 Introduction

Human-computer interaction has not changed fundamentally for nearly twenty years. Conventional GUI techniques appear to be ill-suited for some of the kinds of interactive platform now starting to emerge, with computing devices having tiny and large displays, recognition-based user interfaces using speech and gestures, and requirements for other facilities such as end-user programming. It is assumed that physical interaction between humans and computation will be less like the current desktop keyboard /mouse /display paradigm and more like the way humans interact with the physical world. That is, what we need are interaction techniques well matched with how people will use computers.

Therefore, perceptual user interfaces (PUI) are recently proposed and investigated [1, 2]. Its essence is grounded in how people interact with each other and with the real world. These PUIs [2] are desired to bring our human capabilities to bear on creating more natural and intuitive interfaces, and characterized by interaction techniques that combine an understanding of natural human capabilities (particularly communication, motor, cognitive, and perceptual skills) with computer I/O devices and machine perception and reasoning.

PUIs are expected to require integration at multi-levels of technologies such as speech and sound recognition and generation, computer vision, graphical animation and visualization, language understanding, touch based sensing and feedback (haptic), learning, and modelling and dialogue management. A summary of interaction modalities between human and computer are well shown in [8].

In the following section, we will present several navigation devices developed in MARS lab. Vision-based pointing and wireless personal access system will be discussed in Section 3, and an integrated interaction environment is described in Section 4. Finally, some testing results and a brief discussion are given.

2. Intuitive Navigation for WebSpace

Basic design parameters for interface development are human movements and natural modalities, so that a user does not need to learn unknown equipment and unknown principles of interaction. Intuition is given by using well-known devices familiar to the common user.

2.1 Walking on Treadmill

Figure 1 Virtual Walk Through Paris

How large is virtual space? Using a treadmill makes the user really walk and keep on running [3]. Navigation speed relies on the status of rolling cylinders driven by the walking movement of the user. Buttons configured on bar-sticks in front of the user controls navigation directions. Some researches proposed complicated and expensive huge omni-directional treadmills [4, 5]. A navigation system based on treadmill fits best as a Spatial Navigator in fitness rooms and in emergency-rescue training, where users are forced to walk or run. In fact the treadmill had been in use as test-training device for Berlin fireman and virtual facility management.

2.2 Flying with the Virtual Balance

Virtual Balance™ [6] is a platform reacting on the body movement of a user while standing on it. It is designed for navigation in large virtual environments. This platform is made of weight sensor discs. The sensors receives changes of weight distribution on the disc and transmit them to an analysis unit, which in turn controls the position and orientation of user's viewpoint in the virtual environment. Minimal weight shift on the platform enables navigation. Stepping forward is moving down and walk, leaning backward is going up and fly.

Figure 2 Navigation with Virtual Balance

The virtual balance is perceived as a surfboard or a magic carpet from eastern fairy tales. It allows intuitive navigation by body movements. Virtual Balance requires additional solutions for interaction, e.g. pointing for selection of data.

It is well suited for navigation scenarios in public spaces, e.g. in a museum.

2.3 Virtual Balance as interface for the CAVE

Using the Virtual Balance as Interface to a single user CAVE avoids the clumsy commonly used user CAVE interface such as wired flying joystick or a wand device (Figure 3). Sitting on the virtual balance navigating the Cave enables for relaxed navigation in knowledge discovery spaces. The wired active stereo view is the only obstacle against free body movement. The ability too meet stereoscopic avatars in life size is an appropriate application for the CAVE and requires more development employing microphones, speech recognition and video-based tracking for pointing gestures.

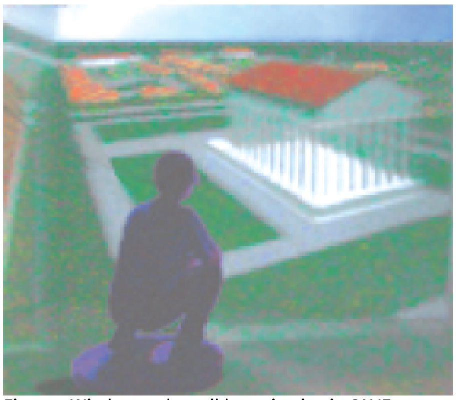

Figure 3 Wireless and tangible navigation in CAVE

2.4 Travelling via Cyberwheel

A series of experiments and testing leads to a 5 degrees of freedom navigation device, entitled Cyberwheel, whose design and implementation is shown in Figure 4. Heading control is performed like on a bike or motorcycle convenient and well known to everyone in principle. Rotation of the handles controls the speed of virtual motion. If the handles are released they will be pushed back to the original position with zero speed. If required, it is also possible to control the pitch angle raising and lowering the upper part of the device. If no more pressure is exerted to change the pitch angle, the pitch control component will return to its original position. Compression and torsion spring are used for force-feedback effect. Rotations of 3 controls change a resistance in rotary potentiometers. The resistances are measured by a micro-controller and sent to a computer via RS-232 cable. Because of its ease and flexibility navigation with the Cyberwheel becomes virtually travelling.

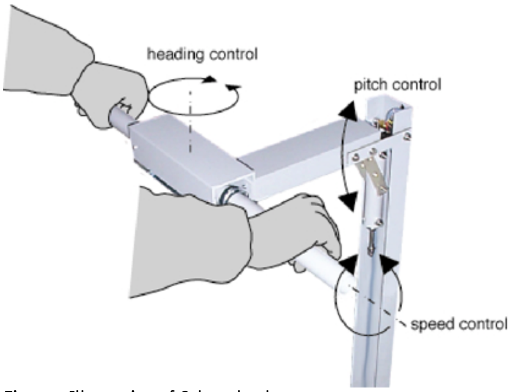

Figure 4 Illustration of Cyberwheel

3. Human-centered non-obtrusive interaction

3.1 Vision-based pointing

We investigated pointing-based interaction by capturing and tracking the arm-movements via computer vision techniques. The role of the vision system is to keep an eye on user's actions, detect a moment when the gesture of pointing has occurred, interpret its direction and follow the pointing hand in real time.

Figure 5 Vision-based Pointing

The system operates in real time, and runs on a Pentium processor under GNU/Linux operation system. In general, the system is able to follow events with 15 frames per second, and can be optimized in different aspects. If the images are digitized at a size of 384x288 pixels, experiments show that the height of the fingertip is approximately 3 Pixels, which result in a maximum resolution of 3.125cm on screen of size 3 by 2 m meter. Virtual screen is back-projected on a wall of approximately 3 by 2 meters, and the user moves freely in a real-world space of approximately 4 by 3 meters. A monochrome infrared-sensitive CCD camera at the ceiling of the space captures the top image of the user with the frequency rate of 25 frames per second. The assumption about the overall lighting is that it has to stay constant during the operation, since colour calibration is done only once during the system initialization. The pointing gesture occurs when the user spreads one of his arms towards the screen. Actually the pointing direction (Figure 5) is restricted to one hemisphere in front of the user, and it is parametrized relative to the projective screen. It is modelled by a 2-D line that connects a point between the user's eyes with the tip of the pointing hand. The recognition of pointing position is then equivalent to finding the horizontal and the vertical coordinates of the point where this line intersects with the screen. A stereo vision system for tracking of pointed (x, y) coordinates will be implemented next.

3.2 Wireless access of personal information

The design principle behind personal wireless access system is to avoid obtrusive typing of various personal data like name, address, personal preferences, and enable user's instant transmission personal data. Ideally, once the user enters the application region, the relevant personal profile will be automatically sent to the application from mobile phone or hand-held computer. Of course, its precondition is that the user has authorized to do so.

We choose Infra-red Data Association (IrDA) standard as the basis to carry out the desired wireless communication of personal data, since it has been widely adopted in the industry, and many IrDA compatible consumer devices, such as notebooks, PDAs, mobile phones, printers, have been available now. Furthermore, IrDA compliant equipment is inter-operative across different applications, manufactures and platforms. The key technical advantages of IrDA standard are:

- simple and low cost implementation,
- low power requirement,
- direct, point-to-point connectivity,
- efficient and reliable data transfer.

In order to implement IrDA-based wireless personal access, we use the infrared adapter supplied by ACTiSYS company. It offers 2.4 meters reliable infra-red link distance under 115.2 Kbps baud rate in most applications using no external power. The adapter is connected to a PC through serial or USB port.

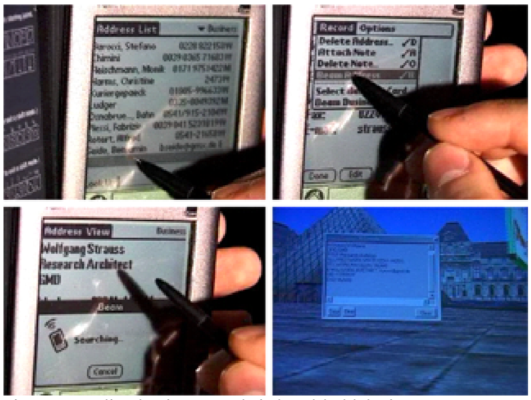

Figure 6 Sending business-card via hand-held device

The file transmitted by an IrDA-enabled device is automatically stored in a specific directory. In order to check whether a new file arrives, the specified directory will continuously retrieved. If a new file is detected, its content will be parsed to get personal information (e.g. user's name or ID), and at last the file itself will be deleted from the directory. A new message, with personal information extracted from the file, will be created and sent to the Application Server (shown in Figure 8). This message may trigger some further actions prescribed by the Application Server (e.g. actually starting the application). A common example (Figure 6) of this personal access is to send a busyness card stored on a user's mobile phone or PDA.

For the time being, Bluetooth technology is beginning to appear on the market, and will make wireless instant transfer of the user's profile even more simple and with better performance. However, IrDA and Bluetooth share the similar communication protocol in application level, and the personal information transferred via Bluetooth can also be stored as file in certain directory in host computer. Therefore, Integration of Bluetooth into current wireless personal access system is easy and simple, requiring little additional work.

4. Integrated Interaction Environment

4.1 Task-oriented Integration

We manage to bring the relevant input/output devices working together, and give a human-computer interface solution for interaction tasks that are commonly involved in e-commerce context. It is assumed that the choice of the input/output device will determine the concrete interaction technique. The prototype installation, which is presented in Figure 7, consists of the following components/devices:

1. Wall size display, it provides a main view of 3D virtual scene.
2. Cyberwheel, it is a 5 degrees of freedom input device as aforementioned, and here is used for easy and precise navigation in 3D environments.
3. Touch screen, it shows current position of the user via displaying a top-view of the 3D scene. In addition it can also be used as application control to perform some operations or choose objects.
4. Infrared receiver, it is for acquiring user's personal information (name, e-mail, preferences) from hand-held infrared-enabled personal devices (mobile phone, Palm Pilot, etc)
5. Video camera, it is reserved for vision-based pointing interaction or for video-conference.

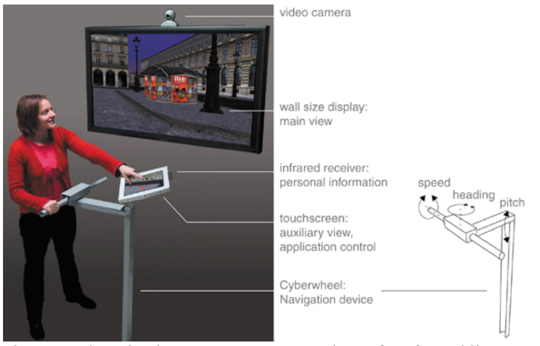

Figure 7 Cyberwheel Prototype - Integrated Interface for Public Space

Its potential e-shopping scenario can be described as follows. When the application is launched, the user may at first be asked to beam via infrared link some personal information, such as name and customer ID, from his mobile device, such as Palm Pilot. Cyberwheel performs the navigation in 3D virtual space projected on the wall-size screen. The user can be aware of his current position from the planar map or the top-view displayed on the touch screen. He may jump immediately to a desired point through a single touch. The selection of virtual objects or performing manipulations (e.g. opening a door or presentation of product) is accomplished in the following way. An icon corresponding to the target object, action or operation will be displayed on the touch screen. The user touches the icon to select the object or trigger the presentation/action/operation. The touch screen serves as somewhat remote control for the virtual environment, and the icon is displayed only when the corresponding action is really applicable. The vision-based pointing technique will permit the user to avoid the use of special pointing device that is necessary to carry in a hand. If the user has made the decision of buying something, the paying-related information such as credit-card ID could be input via hand-held mobile devices.

4.2 Software Platform for Integration

For integration of interface devices, we choose the server/client architecture, in which central server coordinates activities of all input and output devices.

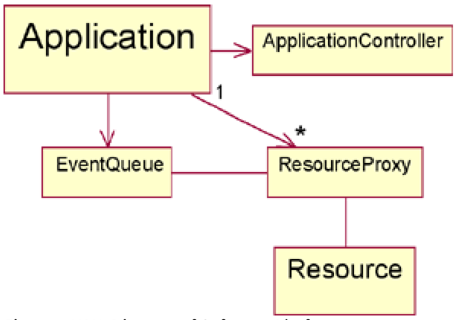

Figure 8 UML Diagram of Software Platform

As shown in Figure 8, all interaction devices send and receive data or requests from application. That is why we do not make clear distinction between them and call them, Resources.

The application server interacts with resources via proxies. It is the task of the proxy to perform connection between application server and the resource itself. With proxy solution even mobile devices may be connected with the application server in a straightforward way.

Application server and a resource interact through communication packets. Thus a resource may be straightforwardly replaced with another one that supports the same functionality. It is transparent for the application server, as each resource sends a data packet through its proxy. When the proxy receives the data packet it may push it directly to the application event queue, or re-direct it for further processing. The application server reads packet from the queue, checks the application status (Application Controller class in Figure 8) and performs alternative action if the data packet does not come as expected. The application server sends command packets to a specific resource possibly with request for the execution confirmation. Finally, the application server redirects the original data packet to the resources interested in accordance with the application scenario in the communication packet.


Figure 9 Integrated Hypermedia Interface setup source/MARS/courtesy of blaxxun interactive

5. Results and Discussion

Perceptual user interfaces will enable multiple styles of interaction via adding human-like perceptual capabilities to the computer, and enhance the computers as tools or appliances, directly enhancing GUI-based applications. Perhaps more importantly, these new technologies will enable broad use of computer assistants, or agents, that will interact in more human-like ways. Towards the goal of setting-up an integrated hypermedia interface (Figure 9), we tested current interaction devices in a series of interactive environments including the following prototype application scenarios:

- Exploration of interactive virtual city. The system can be installed in a travel agency or in an exhibition with a model of a well-known tourist centre. It can be also installed in a museum for exploration of a reconstructed model of an ancient city.
- Interactive 3D product or company presentation. The system can be installed in an exhibition or company reception area. The user walks through a virtual office or shop enquiring information about company products.

These evaluations show encouraging results. The proposed installation works well in these testing scenarios, and basically achieves the desired goals. In the programming level, the lessons we learned from the integration are that: to ease the development of more applications with natural interfaces, we must be able to handle other forms of input (audio, video, ink, and sensor input) as easily as keyboard and mouse input.

In the future there will be an increasing diversity of user interfaces on computerized devices, including various kinds of desk and wall-size computers, as well as devices in everyday objects. Interfaces on these computing devices cannot typically use the standard desktop model. They will play more and more important roles in human-computer interaction.

Acknowledgement

The work in this paper is supported by European Commission Project, SoNG (The Portals of Next Generation, IST-1999-10192).

Reference:

[1] Alex Pentland, Perceptual Intelligence, in: Communication of ACM, March 2000, Vol 43, No. 3, 35-44

[2] Matthew Turk, George Robertson, Perceptual User Interface, in: Communication of ACM, Vol. 43 No 3, 33-34, 2000.

[3] Fleischmann, M., Imagination Systems – Interface Design for Science Art & Education, in: Proceedings IMAGINA', Monte Carlo 1996.

[4] Iwata,H., Walking About Virtual Environments on Infinite Floor: Proc of IEEE 1999 Virtual Reality Annual International Symposium, 1999

[5] Rudolph P. Darken, William R. Cockayne, David Carmein, The Omni-Directional Treadmill: A Locomotion Device for Virtual Worlds. ACM Symposium on User Interface Software and Technology 1997: 213-221

[6] Fleischmann, M.; Strauss, W., Elistratov, V., Interfacing Cultural Heritage, in: Proceedings Eurographics, Milano 1999.

[7] Kolesnik, M., Paar G., Bauer A., Ulm M., Algorithmic Solution for Autonomous Vision-Based Off-Road Navigation, in: Proc. of SPIE: 1998, Orlando, Florida.

[8] Sharma, R., Pavlovic, V. I. & Huang, T. S., Toward multi-modal human-computer interface. Proceedings of IEEE, 1998, 86, 853-86