Real³ Communication and Aromatic Group Computing: HCI and CSCW Research at Canon Media Technology Laboratory

Hideyuki Tamura and Yuichi Bannai

Media Technology Laboratory, Canon Inc.
890-12, Kashimada, Saiwai-ku, Kawasaki 211, Japan
E-mail:{tamura, bannai}@cis.canon.co.jp

INTRODUCTION AND BACKGROUND

The Media Technology Laboratory is one of Canon Inc.'s corporate research labs. Originally called the Information Systems Research Center, the laboratory changed its name when it started research and development in information media. Now, approximately 70 research scientists and engineers, some of them managers, are in charge of research and development mainly in HCI, CSCW, and other fields. Their range of expertise covers AI, natural language understanding, computer vision and graphics, computer architecture, and system software (OS, database, etc.).

REAL³ COMMUNICATION

The basic spirit behind our research endeavors can be summed up by our slogan: "REAL³" (read as "real cube"). That is, our laboratory has three research directions, each representing a type of "reality" of HCI (see Fig. 1).

Figure 1: Real³ Communication

The first "reality" is verbal communication reality in HCI. In this field, we are aiming at "natural" and "smooth"

interfaces for human beings through research into a verbal communication system using a natural language (Japanese). The key technologies in this field are speech recognition and synthesis, and discourse understanding [1]. The focus of research currently centers on noise-durable speech recognition, natural prosody for synthesized voice, and an interface agent which adaptively alters its communication mode according to the user's social position and level of skillfulness.

The second "reality" is visual communication reality, which leads us into the research of virtual reality (VR). In this field, we are studying how to increase photo-reality to improve better human sensations of presence and immersiveness. Our VR system, called "PreView" (Putting Real Environment into Virtual Interactive Electronic World), basically consists of two parts: cyberspace creator and player. For the cyberspace creator, we are developing "mixed reality" technology that amalgamates computer-generated images and real pictures to build cyberspaces more easily and with a certain amount of photo-reality. Recently, we have developed a novel technique to produce a viewpoint-dependent stereoscopic display from multiple images [2]. In this field, we are also investigating compression of 3D images and protocols for 3D communication.

The third type of "reality" is a computer architecture to handle realtime media, — a computer environment that integrates continuous media such as audio and video while maintaining time correlation between real continuous media. Our "optical bus cluster" technology can directly interconnect computers with an optical bus and handle closely-coupled computer resources as if they were a single unit [3]. The computer integration technologies such as parallel architecture and distributed OS form the basis of realtime multimedia communication.

AROMATIC GROUP COMPUTING

"Aromatic group computing" is another term used to represent a field of research in CSCW. This oddly piquant term is based on the following concept. HCI, or CHI,

©1996 by Canon Inc.

3-30-2, Shimomaruko, Ohta-ku, Tokyo 146 JAPAN

© Copyright on this material is held by the author(s)

represents an interaction between a computer and a human being. When two such computers are connected, the humans interacting with the computers can communicate through them. This relation can be expressed as H-C-C-H. The morphology of collaboration between human beings using a multipoint computer network is much like an organic compound in which many Cs and Hs are connected, as shown in Fig 2. This type of computer network may grow by inter-networking, just as flexible combinations of Benzene rings produce various chemical compounds. In the new environments that develop, new approaches must be taken in human-computer and human-human interactions. The term "aromatic group computing" was chosen for this reason.

Figure 2: Aromatic Group Computing

The main target of our CSCW research is to support synchronous cooperative works using high-speed networks and investigate their human factors, — an endeavor to advance the development of what we call "realtime groupware." The themes of actual research include telecollaboration on design work, desktop video conferencing, teleseminar, and remote awareness.

Our research environment for the realtime groupware adopts FDDI (ATM-LAN will be added soon) as a local area network, and an optical fiber network using TCP/IP over ATM protocol for a wide area network. The WAN is derived from participation in the "Joint Utilization Tests of Multimedia Communications" promoted by NTT, and it connects our LANs in remote sites.

HIGHLIGHTS OF RESEARCH ACTIVITIES

Our main HCI and CSCW research activities are software development and evaluation of software usage. Our approach includes the development of our own original interactive devices.

A. Ubiquitous Cameras and Active Awareness

Although media space and video conferencing have been studied by various groups, almost all of the existing systems use fixed cameras to capture their data. In Canon, however, we adopt a specially developed video camera equipped with computer-controlled panning, tilting and zooming functions. Many cameras are now installed for our interoffice and intraoffice communications. Since a user can actively control this system and obtain a remote view, we decided to call it "active awareness." We are currently evaluating the affordance and socialness of this system through its actual use in our office.

B. Networked Collaboration Desk/Pad

Telecollaboration on a high-speed network is expected to make a great contribution in the CAD field. For use in this field, we have developed a workstation with a large flat display (15" or 21") and an ultrasonic digitizer incorporated on top as a pen input device. Users of this workstation can operate line drawings and graphics sharing WYSIWIS display images with others. The flat display is embedded into a desk and its surface can be tilted for easier operation. Our HCI research also investigates human factors in this kind of telecollaboration. We plan to design a pad-type display-unit which can be detached from the desk.

C. Gaze Detection and Control

In the cyberspace player described before, we are studying a gaze utilization method for interaction with 3D space. In Canon, we have already adopted an eye-control mechanism for 35 mm still cameras and video camcorders that enables auto-focusing on the point where the user is gazing. By integrating a detector into a pair of LC shutterglasses or HMD device, the same technology can be used to incorporate the direction of the user's gaze as a key function in HCI. We have already implemented a method to greatly decrease the total rendering load. In this implementation, only the foveated area where a user is focusing is depicted in detail; computer resources are not spent to render detail in peripheral areas. We are also studying a method of identifying 3D objects that are being focused on, and a method of making blinks of the eyes and other user actions into communication commands.

REFERENCES

- [1] Sakai, K. et al., "Robust discourse processing considering misrecognition spoken dialog system", in Proc. Int. Conf. on Spoken Language Processing '94, pp. 895 - 898 (1994).
- [2] Katayama, A. et al., "A viewpoint-dependent stereoscopic display using interpolation of multi-viewpoint images", in Proc. SPIE, vol. 2409, pp.11 - 20 (1995).
- [3] Shibayama, S. et al., "An optical bus cluster system with a deferred cache consistency protocol", in Proc. Int. Conf. on Parallel and Distributed Systems '96 (1996).