Cyber Assist Project for Situated Human Support

Hideyuki Nakashima

Cyber Assist Research Center
National Institute of Advanced Industrial Science and Technology
41-6, Aomi, Koto-ku, Tokyo 135-0064
h.nakashima@aist.go.jp

Abstract

This paper introduces some of the projects conducted at the Advanced Institute of Science in Japan strongly related to the field of distributed multimedia systems. Among several national strategies of AIST, a main target is to create technologies that contribute to make society keep up with this information age so that people will feel safe to live. At the same time, we are cultivating a new industrial field. Specifically, we research and develop a human-centered information support system that is safe and easy to use for anyone. To combine device design, information sensor networks and intellectual agents organically, and to review them thoroughly, including it's information communication infrastructure. Our aim is to make a center, which is able to dispatch this new image of research to all over the world.

Introduction

The current information processing tools such as personal computers and Internet are not always easy to use. Novice users often have to take a class to master them. A new trend in research was initiated recently to turn the table around. MIT's Oxygen project [1,2], Microsoft's Easy Living [3] and Hewlett Packard's Cool Town [4] as well as our Cyber Assist project are among those announcing their new directions: making information processing machines invisible from human users and yet providing rich ubiquitous supporting environment.

The goal of Oxygen project, for instance, is to realize human collaboration aid through the information-processing environment described by the following keywords:

1. pervasive (ubiquitous) – Information processing and communication devices are everywhere.
2. embedded – They sense and modify the real world.
3. nomadic – Users and computation moves around.
4. eternal – Computation does not stop.

Similarly, the research theme of the Cyber Assist Research Center is the development of human-centered information processing technologies, which can provide situated information that *I-want-here-now* through "natural interface"[5]. In other words, we are strengthening a variety of technologies that link digital realm represented by Internet to us people who live in the real world.

The aim of this paper is to introduce our research plans together with our view of the future information-processing environment.

Cyber Assist

Our goal is to develop human-centered information processing assistance systems (intelligence booster) usable without special knowledge or training. We also address the problems of information overload and privacy.

Our target is to propose a plan of the future cities with information feedback control systems. It is achieved through sensors, actuators and information processing over them. Therefore, our use of "cyber" differs from those used by mass media where the word is synonymous to "digital". In fact, we define

cyber = digital + real

meaning that cyber world implies *grounding* of digital, or logical, information to the real, or physical, world.

Our project is classified as follows. We have two main targets:

1. situated information support
2. privacy protection

These targets often contradict each other. When you submit more personal information, like your preference and previous experience, you have better chance of getting personally tailored services. The key issue is that the users getting the services must have the control of their personal information, not the service provider. In this respect, we believe it important that the communication itself is anonymous, and personal identity, if necessary, must be given as the content of the communication.

To achieve the above goals, we have two main approaches:

1. intelligent Content
2. location-based communication

Communication method and the Content are closely related each other and should not be separately designed. We

believe one of the important grounding of Content is to the location (of its existence, use, *etc*.).

Location-Based Communication

One of our main approaches is the development of new communication methodology based on *location*, which is the integration/amalgamation of information processing technology and the real physical world through situated communication.

We are aware that location-based service itself can be implemented without using location-based communication. In fact, mobile phone and GPS can supply similar service. However, GPS cannot be used inside buildings. Furthermore, they are overkill in both infrastructure and power necessary for providing the service in most cases. For example, you may able to buy a can of drink from a vending machine in front of you by using a mobile phone and knowing the location of the phone and the machine. But it is much simpler and straightforward that your hand-held device and the vending machine communicate directly. The identification of both of the machines, such as IP address or a phone number, is not necessary if they use short-range directed communication method such as an infrared beam. We call this location-based communication (in contrast to location-based services). Note also that it is a non-ID communication, in the sense that it does not require ID for addressing. Physical location is used as the address.

There are two possible implementations of location-based addressing, a physical method and a logical method.

The physical method is to focus communication media to a certain physical location. For example, we can direct a laser beam to the target. Communication is limited in the line of sight, but there is less chance of wire tapping or eavesdropping.

The logical method is to use the representation of location as the address of broadcasting (either by radio or sound). This method is usable in conjunction with IP addressing. For example, e-mail may be delivered to the following address:

< x, y, z, t>@room.building.town.country.

We assume that there is an Internet connection to *building* in *town*. The coordinate *x, y, z* is relative to *room*. Alternatively, we can use a global position like

< longitude, latitude, t >@gps.

The time index *t* is important since the target moves around.

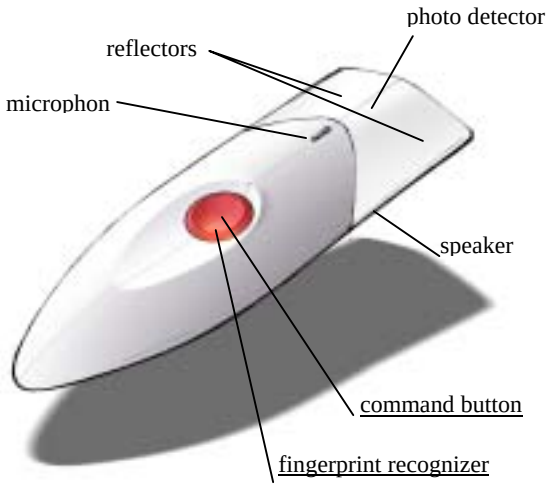

Fig 1 A design of My-button

In both methods, tracking of a target is important to maintain connection. After issuing some request to the server, a device may change its location. Therefore, once a connection is set up, it is necessary that either the server keeps locating the device, or the device keeps reporting its location. An alternative to this method is to set up a temporal ID between the server and the target and use it during the communication (only for logical method).

Being tracked increases the possibility of revealing the user's identity. A user can cut off communication whenever desired, simply by giving a command to the communication device, or by programming it to do so occasionally. As you see in the above example, we have no intention to deny ID-based communication used in Internet. We simply want to extend it.

There are twp positioning methods:

1. Target positioning. The server measures the position of the user. Radar and sonar are among this method.
2. Self-positioning. The user positions itself by the help of infrastructure. GPS is the typical example.

When you move around in a town, and use your tentative location as the address of receiving services, you can keep your identity unrevealed. (Of course, this does not work if you use your permanent location, like your home, as the address.)

Non-ID communication is one of our main proposals and is discussed in detail in this paper. Semantic structuring of information, indexed with location is also our main target.

My-Button

My-Button is a nickname given to our target hand-held or wearable device. As the name suggests, the ultimate design of the device has only one button as the interface, no screen nor keyboard. Just one (or two) click(s) on the button in a proper timing triggers the desired service. The key is that My-Button stays with the user long enough to learn the behavioral pattern of the owner.

The design of my-button reflects our design principle: Keep the channel between human and machine as narrow as possible while keeping the channel among machines or environment maximum.

My-button is equipped with gyroscopes and radio/optical communication media as well as auditory input and output (Fig 1). It communicates with information processing devices in the environment to know its own position and situation. It then keeps track of the history of its usage by the owner and learns his/her intentions contextualized by the situation.

Learning is achieved in cooperation with the environment, since the computing and storage capabilities of the device are limited. Half of the user's history is stored in the environment and retrieved by the key stored in the device. It is important that the information in the environment is meaningful only relative to the information in the device. Even if the servers have full access to their storage, personal information there is either meaningless or unreadable without the help of the personal device.

Since we use location-based communication protocol, the server has no clue to know the identity of the owner of the device. This is the key point of our implementation. Otherwise, our personal history will be open to the public and may be abused.

Fig 2 CoBIT

We have implemented the first version of My-Button as a very simple compact batteryless information terminal called CoBIT (Fig 2)[6]. A CoBIT terminal consists of an earphone, a solar cell and a reflective sheet. In our design, devices embedded in the environment (Fig 3) play a lot of important roles. CoBIT is equipped with a reflective sheet, whose location can be easily identified from cameras mounted in the environment. Then an infrared beam is projected toward the direction from a LED scanner that is capable of tracking the movement of a CoBIT. The infrared signal is modulated using a sound wave and CoBIT replays it as auditory information. The infrared beam itself is the power source of CoBIT, thus there is no need to have any internal power supply.

Cameras in the environment will see the reflection from the reflective sheet of CoBIT to know its location, direction and the movement (gesture). As the total system, it is thus capable of the following:

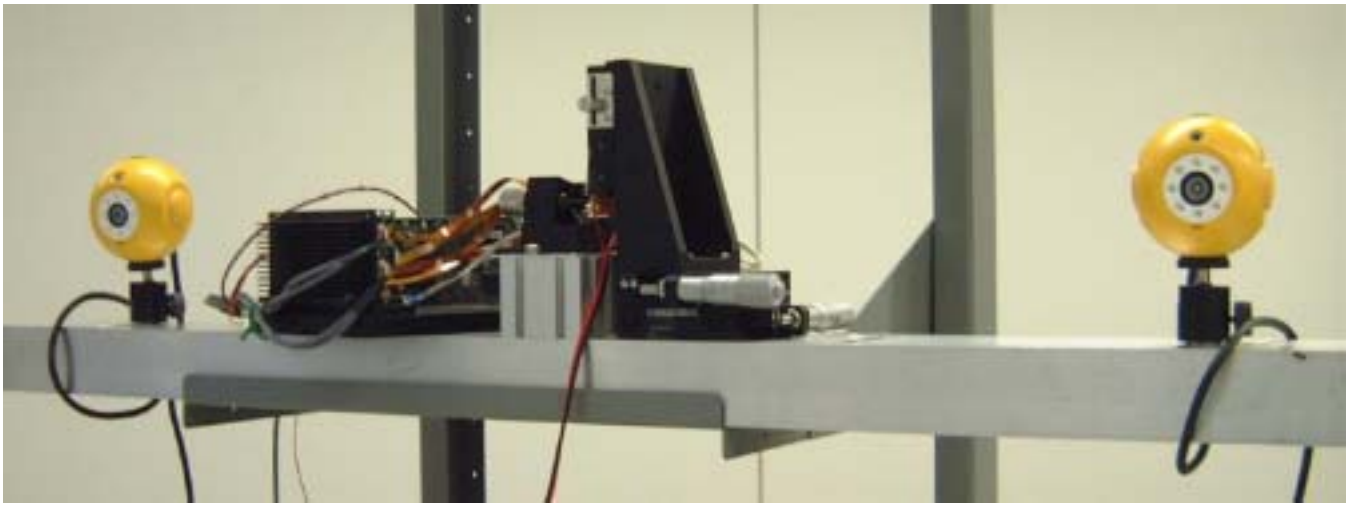

Fig 3 Stereo cameras (both sides) and LED beam scanner (in the center)

1. Downloading information in the form of sound (mainly voice).

2.Uploading information as the movement of CoBIT itself.

We are planning to enforce the upload capability of CoBIT by connecting some short-range radio devices in the near future. In this case, CoBIT will be able to upload its ID and/or other information such as voice stream.

Intelligent Content

A major cause of information overload and digital divide is the semantic gap between humans and computers; humans are better at dealing with deep meaning and machines are better at processing explicit syntax. The only feasible way to systematically fill this gap at the state of the art for the coming several decades or perhaps this century is to make the semantic structure of information content explicit so that machines can deal with them too.

Intelligent content [7] is information content annotated in terms of explicit semantic structure, which is typically predicate-argument structure or semantic network addressing the propositional content

Under location-based communication, the location contributes to plugging the private and the public contents. Namely, your spatiotemporal location allows you to use deixes: linguistic expressions, such as *I*, *you*, *here*, *there*, *now*, *a while ago*, and so forth, which make sense relative to the addressor's spatiotemporal or logical location. Your queries and private content may contain a lot of deixes, and the spatiotemporal location obtained through location-based communication provides the anchor point in the real world that renders those deixes meaningful.

Situated Dialogue

Since the most natural communication media for human is voice and/or vision, and since my-button's primary communication form is sound, natural language dialogue interface is in order. We are planning to develop a situated natural language dialogue system. The simplest dialog may be that the system asks some question to the user and user provides a simple answer and vice versa.

We are developing a system capable of handling a sequence of communications preserving previously referred context. The most difficult part is to detect the context shift. It is not always obvious to determine the part of context that remains valid and the part that should be replaced with a new one.

One of the typical applications of situated dialogue is human navigation. Since the system knows the location (and direction) of the user, deixes are naturally used. If a user has some pointing device (or, my-button can be used as such), even "*this*" or "*that*" are usable.

Application Images

Exhibition Guide

The first version of My-Button, CoBIT, is ideal for low-cost guidance terminal for museum exhibitions. It can provide location-dependent as well as personalized information, possibly tuned to the knowledge level and/or linguistic preference of the guest.

The information content must also be provided in highly indexed way to provide individually tuned explanation of the exhibits. For example, if the visitor is a member of the museum and a frequent repeater, it is possible to skip previously introduced Content and begin with a new one every time he or she returns to the exhibit. We are planning to provide an authoring tool that helps the museum to write such Content as well as its search engine based on semantic matching.

The most challenging part of this application is to provide the feedback cycle to the Content. If, for example, some visitor asks a question whose answer is not provided in the information Content of the museum, there may be some web search (again, it should be semantic, not keyword matching, search) conducted. The new information gathered should be integrated into the content of the museum in a manner it will be available next time.

Cyber Mart

Shops with display of merchandises are just like museums in the sense that giving detailed information of goods to each shopper is important. One can inquire the place of production of vegetables, or the weather condition during their production and so on.

Furthermore, if a shopping cart and stock shelves can communicate each other, there is a large room for supporting shoppers. One of our targets is an automatic payment at the cash register, or shopping without going through the checkout counter.

The route information of the shopping cart (note that the identity of the shopper is not necessary) and timing of buying goods provide the shop owner important data on where and how to display goods.

Information Infrastructure for Rescue

One of an important application of My-button is disaster rescue. My-button's become a secondary information infrastructure, which can form an ad-hoc network using their short-range communications. We are building a

realistic simulator to plan and test the best networking strategies.

The current CoBIT is a passive device and does not function without existence of the infrastructure such as infrared emitter. Since we may not be able to rely on the infrastructure in case of large disaster, we need more autonomous devices. The next model of CoBIT will be coupled with radio communication devices and this model is our target for rescue.

Summary

The current information processing technology is mainly applied in the area where expert users are the target. With the advance of communication technology and its wide-availability, it is now time to apply it to help ordinary people in their everyday life. Many proposals are made in the direction, assuming the realization of ubiquitous information environment. Cyber Assist research project is one of them.

In Cyber Assist, we focus on location-based human support in Content and communication. We presented the concept of optical location-based communication together with some of its application images.

Acknowledgment

Most of the content of this paper was made concrete thorough research activities of the members of Cyber Assist Research Center.

References

[1] Michael L. Dertouzos: *The Unfinished Revolution.* HarperCollins Publishers, Inc. 2001

[2] http://www.oxygen.lcs.mit.edu/

[3] http://www.research.microsoft.com/easyliving/

[4] http://cooltown.hp.com/cooltownhome/index.asp

[5] Hideyuki Nakashima and Koiti Hasida. Location-based Communication Infrastructure for Situated Human Support. In *Proceedings of SCI 2001(World Multiconference on Systemics, Cybernetics and Informatics)*, 2001.

[6] Takuichi Nishimura, Hideo Itoh, Yoshinov Yamamoto and Hideyuki Nakashima. A Compact Battery-less Information Terminal (CoBIT) for Location-based Support Systems, In *Proceedings of International Symposium on Optical Science and Technology (SPIE)*, 4863B-12. 2002.

[7] Koiti Hasida. GDA: Annotated Document as Intelligent Content. Invited talk at *COLING 2000 Workshop on Semantic Annotation and Intelligent Content*, Luxembourg. 2000.