Augment-able Reality:

Situated Communication through Physical and Digital Spaces

Jun Rekimoto

Sony Computer Science Laboratory

3-14-13

Higashigotanda, Shinagawa-ku

Tokyo 141-0022 Japan

rekimoto@csl.sony.co.jp

Yuji Ayatsuka

Information Technology Laboratories

Sony Corporation

Kitashinagawa, Shinagawa-ku

Tokyo 141-0001 Japan

aya@csl.sony.co.jp

Kazuteru Hayashi

Dept. of Information Science

Tokyo Institute of Technology

Ookayama Meguro-ku

Tokyo Japan

k-hayasi@is.titech.ac.jp

Abstract

Most existing augmented reality systems only provide a method for browsing information that is situated in the real world context. This paper describes a system that allows users to dynamically attach newly created digital information such as voice notes or photographs to the physical environment, through wearable computers as well as normal computers. Attached data is stored with contextual tags such as location IDs and object IDs that are obtained by wearable sensors, so the same or other wearable users can notice them when they come to the same context. Similar to the role that Post-it notes play in community messaging, we expect our proposed method to be a fundamental communication platform when wearable computers become commonplace.

1 Introduction

Augmented Reality (AR) systems are designed to provide an enhanced view of the real world through see-through head-mounted displays[16] or hand-held devices[11]. Various kinds of context sensing technologies, such as position sensors[5, 4], or ID readers[11], are used to determine digital information according to the user's current physical context. Compared with traditional information retrieval systems, users of AR systems can get context-aware information without being bothered by cumbersome operations that might impede the user's real-world tasks.

The concept of wearable computing is pushing this direction even further. Wearable computers are always ''on'' always acting, and always sensing the surrounding environment to offer a better interface to the real world.

Taking into account recent advances in wearable devices, we can expect that people will soon carry their wearables at all times, and that our lives will be constantly supported by context-sensitive information from those wearables.

While most of the existing augmented reality systems are mainly focusing on context-aware information presentation, information registration interfaces for AR are yet to be investigated. For example, KARMA[4], which is a well-known AR system, displays information about the laser printer based on the current physical position of a head-worn display. Such information is assumed to be prepared beforehand. To add new information during a task, the user has to return to the normal computer environment. In that sense, current AR systems are essentially ''context-sensitive browsers'' for the real world, and information is limited to a one-way flow. Since two-way communication tools (e.g, E-mail) have fundamentally different functionality and applications from one-way tools (e.g., radio), we can imagine several new usages of AR if it allows bidirectional communications.

This paper presents an environment that supports information registration in the real world contexts through wearable and traditional computers. Using this technique, a user of wearable computers can dynamically create a data item such as a voice or a photograph and can attach it to nearby physical objects or locations. To enable this, the system supports drag-and-drop operations between personal and context-sensitive spaces. The user of traditional computers can also add data to the physical environment through the Web and E-Mail interfaces. By combining these information flows, we can freely transfer data from physical space to digital space, or vice versa. We call such an environment "augment-*able* reality", because users of our system are not just browsing information situated in
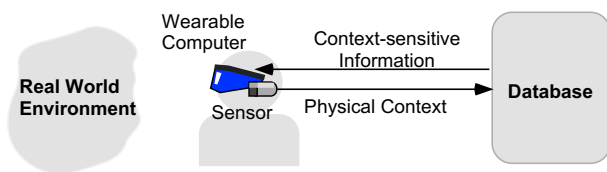
(a) Traditional Augmented Reality Systems
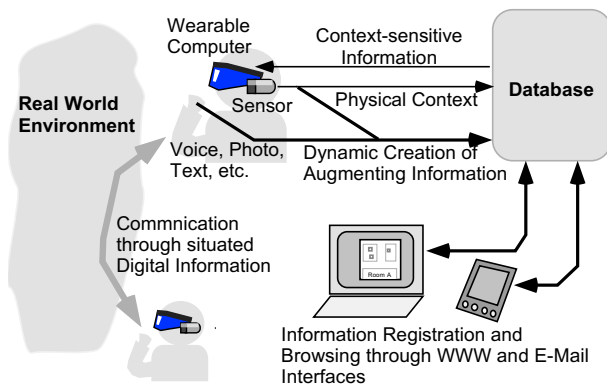
(b) Augmentable Reality Environments

Figure 1: The Augment-able Reality Concept

the real world, but can also create, attach, or carry the data with them.

2 Augment-able Environments

Let us start with a scenario when augment-*able* reality becomes commonplace.

When you enter the video studio, your eyeglass notifies you that a new video effector has been installed. You approach the effector and find a movie that was attached by your colleague, describing typical usage of equipment. Meanwhile, you find the VCR is broken and the tape has been damaged. You create a voice-note with a still picture of the damaged tape and attach it to the VCR for other users.

When you go outside for lunch, you find several virtual messages are floating in front of restaurants. Some are commercials, but many are messages that were created and attached by previous visitors, giving information and ratings on the restaurant. You select a restaurant by looking at these messages. While you are at lunch, you remember that you will have visitors at noon. You select an icon representing your office in the eyeglass, and attach a voice memo to it ....

As we can see in this example, the concept of augment-able reality can be summarized in features outlined in the following sections.

Dynamic creation of augmenting information

People wearing computers and having network facilities can dynamically create information and attach it to the user's surrounding physical environment. From the system's point of view, this is achieved by storing a created data item with contextual information obtained from several wearable sensors. Our current prototype supports infrared beacons and visual ID markers as contextual information. Another possibility is to use geographical locations based on GPS (global positioning system) or PHS (personal handy phone) as contextual information.

There are several possibilities about data types that can be attached to the environment. Our current prototype supports voice notes and photograph snapshots. Other possible data types are text (created by a wearable keyboard) or hand-drawn strokes (created by a miniature touch panel).

Context-sensitive information notification

Similar to other AR systems, the attached data can be browsed by (the same or other) users who are wearing computers. The concept of ''wearing'' is essential; otherwise users would not notice digitally attached information.

Information sharing among wearable and normal computer environments

Attached information can also be accessible from users in normal computing environments. For example, when a wearable user attaches a voice note to the meeting room, it becomes visible on the floor map page of the Web browser. Conversely, normal computer users can attach data to the physical context through the Web or E-Mail interfaces, and wearable users will notice the data when they come to the corresponding situation.

Communication through situated information

The combination of these features creates a new way of communication. Instead of sending messages to people, we can place message in "physical contexts" to indirectly communicate with other people. After having dinner at a restaurant, for example, people might leave their impressions or ratings of that restaurant at that location. Afterwards, other people would be able to check the
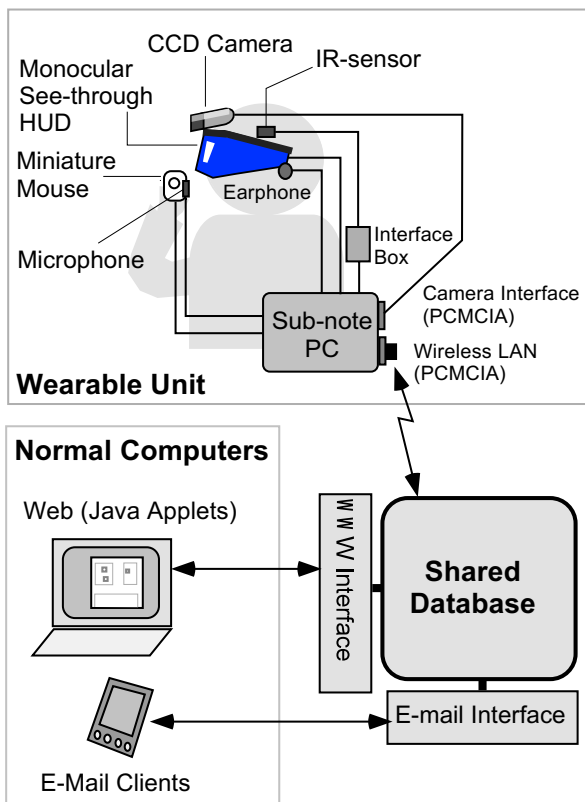
Figure 2: Schematic of the prototype system

Figure 3: Two ID systems deployed in the environment (left: infrared beacons, right: printed 2D matrix codes)

messages from the front of the restaurant, or when they are browsing a map on the Web.

Even when messages are directed to the person, we can select several contextual attributes to do that. For example, you may want to attach an e-mail message to a person's office door (locally or remotely through the net), to get their attention when they return to the office.

3   System Design

To achieve the proposed concept described in the previous section, we are currently developing a set of prototype systems based on wearable and normal desktop computers. Figure 2 shows the overall system design of the prototype. Systems share the same database server on the network through wired or wireless communication. If one user attaches a data item to the particular physical context (e.g., location), this effect immediately becomes visible from other computers.

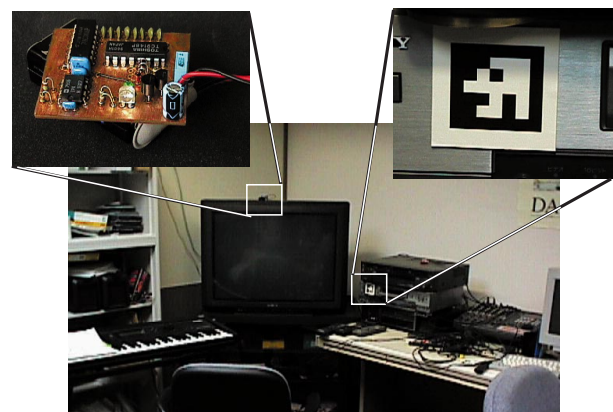The following subsections present details of each subsystem as well as their user interfaces.

3.1   Environmental Support

To make it easier for wearable computers to capture surrounding contexts, we have deployed two kinds ID systems in the environment. One is an infrared (IR) beacon that periodically emits a unique number to the environment (Figure 3, left); the other is a printed 2D matrix code that can be recognized by a head-worn camera (Figure 3, right).

One noticeable difference between these two ID systems is emission range. While IR beacons can cover room-size areas and are relatively robust regarding orientation of the sensors, printed IDs are more sensitive to distance and the camera orientation. We are thus using IR beacons to detect current rooms (e.g., meeting room, office, etc.) and printed 2D codes for object level identification (e.g., VCR, bookshelf, etc.).

Since 2D codes are virtually costless and printable as well, there are some usages that could not be achieved by other ID systems. For example, we can use small Post-it cards with a 2D code. This (physical) card can convey digital data such as voice notes or photographs with attachment of the ID.

3.2   Wearable System

Hardware

The head-worn part of the wearable system (Figure 4, above) consists of a monocular see-through head-up display (based on Sony Glasstron), a CCD camera, an infrared sensor (based on a remote commander chip for consumer electronics). These devices are connected to a sub-note PC (Mitsubishi AMiTY) communicating to the network through a spread-spectrum wireless LAN (Netwave Air-
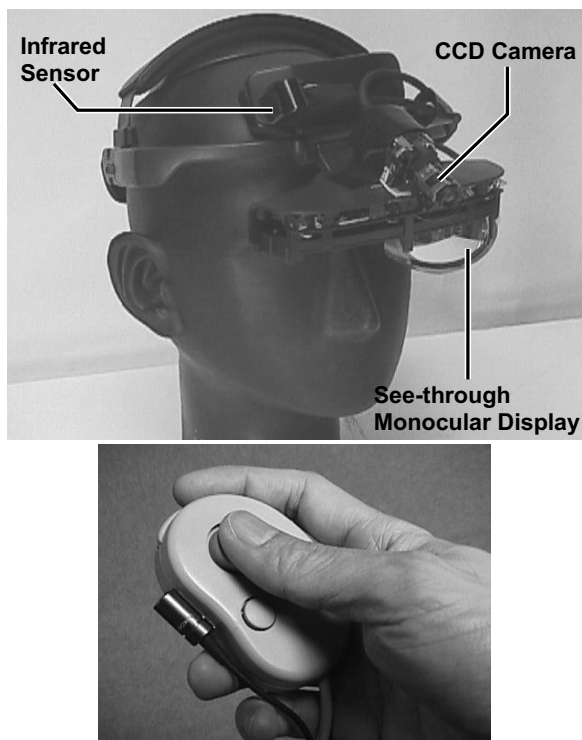
Figure 4: The head-up and hand-held parts of the wearable system

Surfer). The user controls the system with a miniature pointing device. A microphone attached to the pointing device is used to make voice notes (Figure 4, below). The head-worn camera is primarily used to detect 2D codes in the environment, but is also used to take still image

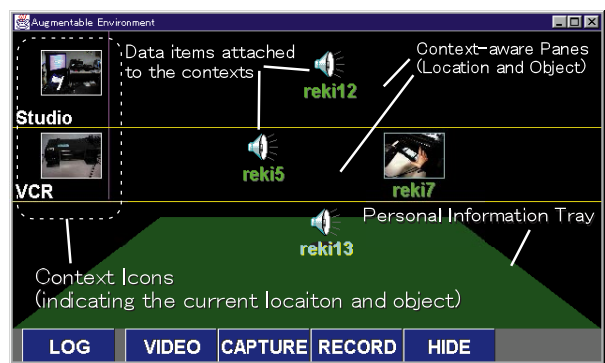Figure 5: Wearable system is used to attach a voice note to the physical object

Figure 6: Screen layout of the wearable display. The user looks at this image on their monocular see-through display. The above two areas (called context-aware panes) are corresponding to the current physical contexts (location and object), while the bottom area (called the personal information tray) stores personal data carried by the user.

pictures. Figure 5 shows a user creating a voicenote with the wearable unit.

User interface

As a context-aware browser for the real world, this wearable system acts mostly the same as our previous AR system NaviCam[11]. The system recognizes the surrounding environment by recognizing attached visual tags on physical objects, or by detecting infrared beacons installed at locations in the environment. For example, when the system recognizes a visual tag attached to the door of the meeting room, the meeting schedule would appear on the see-through head mounted display. In a same way, when the user is walking into the cafeteria, an infrared sensor automatically detects the location by receiving an IR ID from that location. Since the system is equipped with a wireless LAN, information can dynamically be retrieved from the (shared) database on the network.

The unique feature of our new system is its ability to add new data to the physical environment. To enable this feature, the user switches the system to ''*authoring mode*.'' A window consisting of a personal information area and context sensitive area then appears on the see-through display (Figure 6). The personal information area (also called "the personal information tray") stores the data carried by the user. The system currently supports voice notes and still images (captured by a head-worn camera) as such data.

The other area, called the ''*context-aware area*,'' shows information related to the current physical context. The

area is further split into two panes, the location pane and the object pane. These panes represent location-level and object-level contexts. The location-level context is detected by the IR beacons, while object-level context is determined by visual tag recognition.
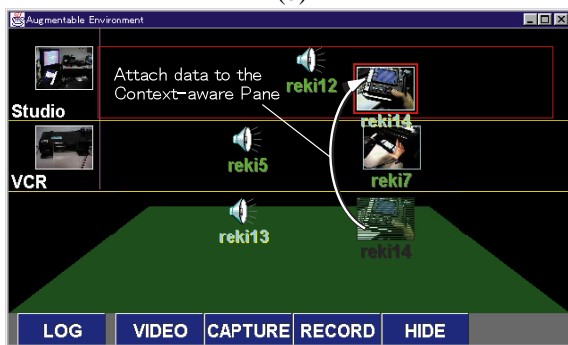
These context-aware panes change when the user moves to a new situation. For example, when a user enters a

(a)
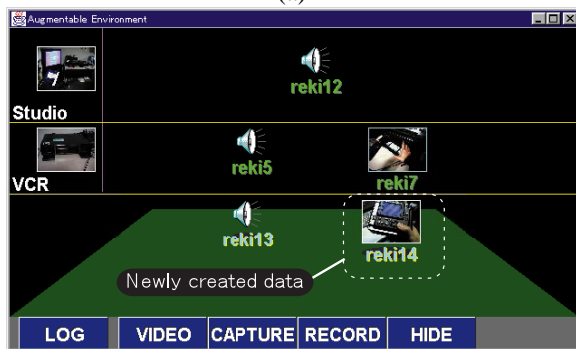
(b)

(c)

Figure 7: Creating and attaching data to the physical context. (a) The user first presses the ''CAPTURE'' button to take a snapshot and adds a voice note. (b) The newly created data object then appears on the user's personal information tray. (c) The user drags it to the context-aware pane to attach the data to the current location (the studio).

meeting room, IR beacons from the room tell the wearable system the meeting room ID, the location pane switches to the meeting room and icons attached to that location appear. A gentle sound also rings to announce this context switch. The user can then open or playback these icons by pressing one of them with the right button of the hand-held mouse.

To support intuitive data transfer between personal and context-aware spaces, the system provides *drag and drop* operations between these areas. That is, the user can create a voice memo, (which will appear on the user's personal area as an voice icon), then drag and drop it to the location pane to attach it to the current recognized location (Figure 7). Since the operation is bidirectional, the user can also drag an object from the context-aware panes to the personal tray to copy information that is attached (by someone or by the user himself) on the current physical context.

For example, suppose that you find that the VCR in front of you has a problem and may damage video tapes. You create a voice note such as ''This VCR is broken, do not insert tapes, or they may be damaged,'' with a microphone while pressing a record button on the miniature pointing device. An icon representing this newly created voice note then appears on the personal tray. You also see that the object-sensitive pane is showing the VCR name and picture to indicate the system is recognizing it. You then drag the voice note icon from the personal tray to the VCR pane; the voice note is now attached to the VCR. Afterwards, other users (wearing a computer) who try to use the VCR will find your warning.

Time Machine Mode

The context-aware panes are not always directly linked to the current physical situation. When the user taps on the left side of the window, the system switches to what we call the ''Time Machine Mode'' (Figure 8).

In this mode, the user can navigate through previously visited rooms or objects without physically visiting these that contexts, by clicking arrow buttons on the context panes. This interface is similar to the forward and backward buttons on web browsers. The user can also visit the *future (past)* by clicking buttons around the time indicator.

Combining these navigation techniques allows a user to attach data to any context at any time. For example, suppose that you want to attach a meeting agenda to the conference room but it should not become visible until next Monday. To do this, you first (virtually) visit the meeting room in the Time Machine Mode, change the time to the
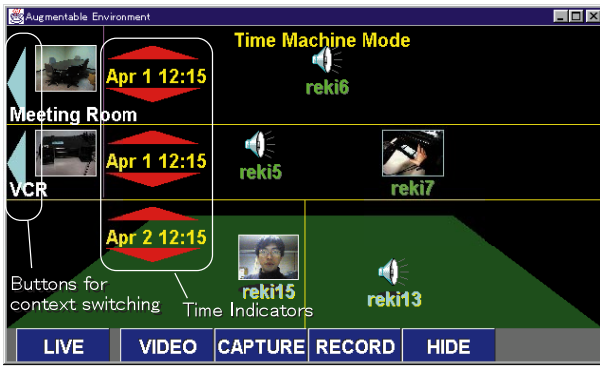
Figure 8: The time machine mode allows a user to change the location and time of the context-aware panes, to access data which are not available in the current physical context.

next Monday, then drag the agenda from your personal tray to the meeting room pane. It is also possible to move data to the future personal tray; it can be used as a personal reminder. To encourage this technique, the personal tray is separated into two parts (left and right) during the Time Machine Mode. Like context-aware panes, the time of the left part can be changed, while the right part always stays with the *current* time.

3.3 Desktop Environments

Web (Java Applet) Interface

Attached information can also be accessed from normal computing environments. We have developed a Java applet for retrieving and adding information to the physical environment (Figure 9). The user can display or playback attached data simply by clicking a corresponding icon on the floor map. The user can also attach data to a physical location or an object through the map window of the applet. This attachment immediately becomes visible to other users. Since this applet can be started from Java-enabled web browsers (such as Netscape or Internet Explorer), the user can access the attached data from virtually anywhere.

E-Mail interface

We also provide an E-mail gateway to the system. For example, when users wish to attach a message to the meeting room, they can simply send mail to that room, such as:
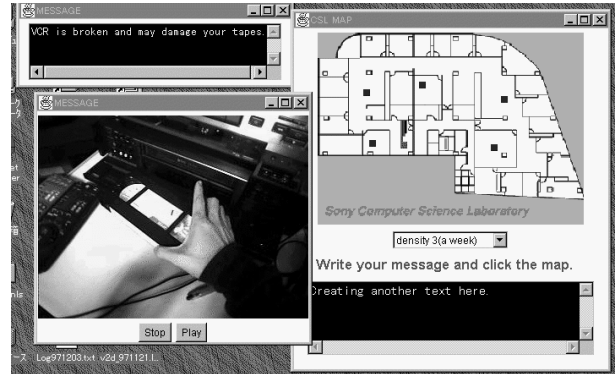

Figure 9: Accessing situated information from a Java applet

```
To: room3@ar.csl.sony.co.jp
Subject:
Msg=Today's meeting is cancelled.
Time=4/6/1998 10:00
Timeout=1 day
```

(Note that "Time" and "Timeout" fields are optional). This capability encourages a user to attach information from a remote and mobile environment (typically using a PDA with a limited communication bandwidth).

4 Discussion

Graceful notification

Based-on our initial experience with the prototype system, we feel that the key design issue on augment-able reality is how the system can gracefully notify situated information. If the system relies on rustic beeps or pager vibration, we would be constantly interrupted by the system. Our current approach is to overlay information on a see-through heads-up display; we expect this approach to be less obtrusive when the display is embedded in eyeglasses [14]. Another (and perhaps more handy) approach would be having a very simple notification indicator such as a combination of small LEDs on eyeglasses. When a user notices the presence of situated data, they can browse it through a palmtop or wristtop display.

Comparison with other approaches

Some part of augment-*able* environments can also be achieved by simply attaching ''Post-It'' notes to the objects. However, we believe that there are several features that can only be available by digitally attaching the data. First,

attached information can also be accessed from another environment, typically a desktop-computing environment. In the same way, a wearable computer user can browse information that was registered remotely from other desktop computers. Secondly, we can apply several information retrieval techniques to filter attached information. Keyword search, relevance matching, additional contextual information (owner, timestamp) are typical candidates for filtering information. For example, users browsing attached data while in a meeting room can filter out data items that do not contain ''*project XYZ*.'' Beyond that, it is also possible to supply special attributes. For example, the user can attach data with attributes such as ''a week later'' or ''every Monday.'' Filtering also solves privacy issues. While Post-it notes are visible to anybody, digital data could be disclosed to selected people only. Third, digital information can convey more information than Post-Its. Voice notes, still images or movies, as well as text messages can be created and attached using wearable devices. Finally, on some occasions, physical attachment is simply impossible. We can digitally attach data to the current global position based on longitude-latitude information from the GPS, but attaching Post-Its to that place is not always possible or practical.

The other alternative is to deploy small displays ubiquitously in the environment. We can attach data to these displays locally or remotely through E-mail to that display. Although some features are only available by the wearable approach (e.g., privacy), it is an interesting open issue on how these two approaches can supplement each other.

5   Related Work

Forget-me-not[8] is a wearable remembrance aid that constantly logs physical context to retrieve information based on the user's personal history. The Wearable Remembrance Agent[12] also helps remind people of things by continuously displaying information relevant to the current physical context. This system also supports storage of newly created information tagged with physical contexts for later retrieval. Both systems are mainly used as personal memory aids, and do not explicitly support *attaching* digital data to the physical environment (which will be shared by other users).

In the context of AR and wearable systems, the idea of deploying machine readable IDs in the environment is gaining popularity. Our current system uses infrared and visual markers that were originally developed for our previous system NaviCam[11]. Similar techniques are also used in other AR systems[2, 9, 15]. Other possibilities include wireless tags[1], near field radio[7] and the global positioning system (GPS).

Some might argue that the ultimate system should not depend on such artificial IDs, probably by employing more sophisticated technologies such as natural scene understanding. We disagree. Since humans rely on artificial markers such as traffic signals or indication panels, we believe that having appropriate artificial supports in the physical environment is the key to a successful AR and wearable systems.

Turing Machine[3] is a prototype wearable system for tourists using a differential GPS and other sensors. The system overlays geographical information such as building names within the user's sight. This system would be more attractive if it allowed the users to leave messages at those locations (instead of scribbling graffiti on historical sites!).

Our interface design for attaching data is partially inspired by ''Fix and Float''[13], an interface technique for carrying data within a virtual 3D environment. ''Pick and Drop''[10] provides a method for carrying digital data within a physical space using a stylus. Augment-able Reality extends this concept, by regarding our body (with a wearable computer) itself as a carrier of digital information. Regarding the ''Time-Machine mode'', a document management system called ''LifeStreams''[6] also supports similar techniques; to make a document a reminder, this system allows a user put a document at a future time of the document stream.

6   Conclusion and Future Directions

In this paper, we have described ''augment-*able* reality'' where people can dynamically create digital data and attach it to the physical context. We have developed working prototypes to explore this concept and has created user interfaces that supports easy data transfer between personal and context-aware information space.

We are also currently working on a system using GPS, based on the same concept. With this system, the user will be able to digitally leave a hand-written note or other data type on the current geographic location, where others can observe the attached data from wearable and traditional computing environments.

Acknowledgments

Kohtake for hardware supports and Tsukasa Yoshimura for his contribution on early exploration of augment-able concepts.

References

[1] Yuji Ayatsuka, Jun Rekimoto, and Satoshi Matsuoka. UbiquitousLinks: Hypermedia links embeded in the real world. In *IPSJ SIGHI Notes*, number 67-4, pp. 23--30. Japan Information Processing Society, July 1996. in Japanese.

[2] Benjamin B. Bederson. Audio augmented reality: A prototype automated tour guide. In *CHI'95 Conference Companion*, pp. 210--211, 1995.

[3] Steven Feiner, Blair MacIntyre, Tobia Hollere, and Anthony Webster. A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. In *Proc. of the first international symposium on wearable computers (ISWC '97)*, 1997.

[4] Steven Feiner, Blair MacIntyre, and Doree Seligmann. Knowledge-based augmented reality. *Communication of the ACM*, Vol. 36, No. 7, pp. 52--62, August 1993.

[5] George W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. *Communication of the ACM*, Vol. 36, No. 7, pp. 38--49, July 1993.

[6] E. Freeman and D. Gelernter. Lifestreams: A storage model for personal data. *ACM SIGMOD Bulletin*, March 1996.

[7] Richard Hull, Philip Neaves, and James Bedford-Roberts. Towards situated computing. In *Proc. of the first international symposium on wearable computers (ISWC '97)*, pp. 146--153, 1997.

[8] Mik Lamming and Mike Flynn. Forget-me-not: Intimate computing in support of human memory. In *FRIEND21 '94 International Symposium on Next Generation Human Interface*, 1994.

[9] Sue Long, Dietmar Aust, Gregory D. Abowd, and Chris Atkeson. Cyberguide: Prototyping context-aware mobile applications. In *CHI'96 proceedings*, 1996.

[10] Jun Rekimoto. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In *Proceedings of UIST'97*, pp. 31--39, October 1997.

[11] Jun Rekimoto and Katashi Nagao. The world through the computer: Computer augmented interaction with real world environments. In *Proceedings of UIST'95*, pp. 29--36, November 1995.

[12] Bradely J. Rhodes. The wearable remembrance agent: a system for augmented memory. In *Proc. of the first international symposium on wearable computers (ISWC '97)*, 1997.

[13] Georg G. Robertson and Stuart K. Card. Fix and Float: object movement by egocentric navigation. In *Proceedings of UIST'97*, pp. 149--150, October 1997.

[14] M. B. Spitzer, N. M. Rensing, R. McClelland, and P. Aquilino. Eyeglass-based systems for wearable computing. In *Proc. of the first international symposium on wearable computers (ISWC '97)*, 1997.

[15] Thad Starner, Steve Mann, Bradley Rhodes, Jeffrely Levine, Jennifer Healey, Dana Kirsch, Rosalind W. Picard, and Alex Pentland. Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments*, No. 4, pp. 386--398, 1997.

[16] Ivan Sutherland. A head-mounted three dimensional display. In *Proceedings of FJCC 1968*, pp. 757--764, 1968.