NAIST-IS-DT0061202

Doctoral Thesis

Design and Implementation of Augmented Memory

Tatsuyuki Kawamura

February 6, 2003

Department of Information Processing
Graduate School of Information Science
Nara Institute of Science and Technology

Doctoral Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
DOCTOR of ENGINEERING

Tatsuyuki Kawamura

Thesis Committee: Masatsugu Kidode, Professor
Shunsuke Uemura, Professor
Hideaki Takeda, Professor (National Institute of Informatics)
Yasuyuki Kono, Associate Professor

Design and Implementation of Augmented Memory*

Tatsuyuki Kawamura

Abstract

This thesis proposes the design concept of a practical computational augmentation of human memory (augmented memory) for everyday life. This thesis also describes the techniques for implementing the concept as well as their experimental results with good performance.

Varieties of practical memory-aid systems are proposed to help a user for both memorizing his/her experiences and recalling a desired them at anytime and anywhere. People undoubtedly wait for systems to computationally augment human memory to be distributed. Actually, a person's memory activities are used not only for recollecting events he/she has experienced but also for solving his/her current concerns, and planning future plans throughout his/her everyday life. However, errors in his/her memory activities, e.g., slip, lapse and mistake, easily occur. Such errors make him/her give up his/her scheduled task. The ultimate goal of the augmented memory is to enable the user to reduce such errors by integrating seamlessly his/her natural human memory with the computational augmented memory in his/her everyday life. This thesis first describes how the augmented memory should be designed. Although the augmented memory has been extensively studied in recent years, most of studies have not presented what the definition of the augmented memory is. This thesis also describes how augmented memory modules should be implemented.

* Doctoral Thesis, Department of Information Processing, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DT0061202, February 6, 2003.

By redefining the conventional concept of Rhodes's augmented memory to my new concept, this thesis discriminates ideal specifications of the augmented memory into "Encoding," "Storage," and "Retrieval" processes. In order to develop the augmented memory, this thesis proposes a framework, which is termed *SARA*, for accomplishing the augmented memory. The following two categories are defined to design the framework. One is an operational category: Memory Retrieval, Memory Transportation, Memory Exchange, and Memory Editing. Another is the associable category: Spatial, Physical (Physical Human and Physical Object), Temporal, Behavioral, and Psychal elements.

This thesis implements three types of augmented memory modules on the above design concept of the augmented memory. All modules employ a wearable camera, which can capture a scene from a user's viewpoint, and a wearable computer. (1) The *Residual Memory* activates a user to recall a "location-triggered" past event. Suppose that the user tries to recall past event he/she had experienced at the place where he/she has ever been there when he/she goes there. In this module, the main topic of this work is a design and an implementation of video retrieval techniques in the real world. This module represents a practical example of the operational category of the "Memory Retrieval" element. This module also shows a practical example of the associable category of the "Spatial" element. In order to design this module, I utilize "Continuity" and "Spatiality" characteristics, which are the transitions caused by person's spatial and temporal movements, in designing this module. This module employs three types of techniques: (i) a stable image matching method excluding user's head motion and moving objects in a scene captured from a wearable camera using two mono-axis gyro sensors, (ii) a video scene segmentation method to detect a video scene comprehensible for the user using a moving average method with two mono-axis gyro sensors, and (iii) the real-time video retrieval method using two different types of feature spaces composed of a time-sequential space and a color image-feature space. (2) The *Nice2CU* activates a user to recall a "human-triggered" past event and information, e.g., name, affiliation, birthday, etc. In this module, the main topic of this work is a design and an implementation of an augmented memory management method using persons' profile data and meeting logs. This module

represents a practical example of the operational category of the "Memory Retrieval" and "Memory Editing." This module also shows a practical example of the associable category of the "Physical Human" element. This module employs a Radio Frequency Identification (RFID) device for registering a target person. This module proposes a "Card and Mirror" interface for a easy registration and an automatic update of the target person's profile. The card interface employs a business card attached to an RFID tag. The mirror interface is composed of a magic mirror and a camera set behind the mirror. This module manages a person's information as "Profiles," "Experiences," "Messages," and "Human Relations." This module employs three types of techniques: (i) the easy registration method of the target person using the business card attached to the RFID tag to reduce the workload for registering the target person, (ii) the automatic update method of the target person's information using the mirror interface to minimize the cost for updating persons' information, and (iii) a rating method of human relations using persons' profile and meeting logs to recommend "Experiences" and "Messages" videos to the user. (3) The *Ubiquitous Memories* activates a user to recall an "object-triggered" past event. Suppose that the user tries to recall past event he/she had experienced from the plaque he/she got when he/she won first prize in the 100-meter dash at an athletic festival. In this module, the main topic of this work is a design and implementation of a module containing cognitive operations. This module represents an example of the operational category of "Memory Retrieval," "Memory Exchange," and "Memory Editing." This module also shows an example of the associable category of the "Physical Object" element. This module employs the RFID device for associating a video captured by a wearable camera with a real world object. Each real world object is attached to an RFID tag. The module is designed by using human cognitive trait of the "Encoding Specificity Principle."

Keywords:

augmented memory, wearable computing, ubiquitous computing, memory-aid in everyday life

Dedication

Acknowledgements

I have been influenced by many people around me during the six years I have been at NAIST. They have shaped the way I do research and introduced me how to study.

I would like to first thank Prof. Masatsugu Kidode for accepting me and my research themes, which is for computational augmentation of human memory in our everyday life. I would like to thank Prof. Shunsuke Uemura. I would like to thank Prof. Hideki Takeda for the valuable advice since I started belonging to the artificial intelligence laboratory. I would like to thank Associate Prof. Yasuyuki Kono for providing me with guidance and support, and always the allowing to try something new and different research themes.

Next I want to thank the augmented memory group of graduate students - Takahiru Ueoka, Satoshi Murata, Yutaka Kiuchi and Chin-Chan Lin. Takahiro Ueoka is the most familiar colleague in the artificial intelligence laboratory members. I have discussed a lot research topics of augmented memory. He has also assisted conducting experiments of a research topic named *Integrated Probabilistic Histogram (IPH)* (Chapter 4) - thank you. Satoshi Murata have assisted implementing one of augmented memory modules termed *Ubiquitous Memories* (Chapter 6) and conducting experiments for the module. I believe that he surely takes over my contributions - thank you. Yutaka Kiuchi has also discussed the *IPH* and assisted conducting experiments of it - thank you. Chin-Chan Lin has belonged to our research group as one of research students and been in the Institute of Electronic Engineering, Southern Taiwan University of Technology. He has been a good assistant for implementation of one of modules termed *Nice2CU* (Chapter 5). He has assisted conducting experiments of the module - thank you.

I would like to thank my parents. Both my father and mother are unconditionally the most important persons in my whole life - thank you so much again.

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Background and Motivation

I first give you the following scene:

Ally: "Did you ever wich that you could erase a whole day from your life?"

Ally: "Or, or, a year, even?"

Cage: "How could you even think of erasing this one?"

Cage: "So much has happend."

Ally: "A lot of it bad."

Cage: "Some good, some bad."

Cage: "Most of it memorable."

Cage: "How many people get that?"

Cage: "My mother used to tell me..."

Ally: "Oh, God."

Ally: "I'm ssory..."

Ally: "What, what did your mother say?"

Cage: "She'd say if you think back and replay your year..."

Cage: **"if it doesn't bring you tears, either of joy or sadness, consider the year wasted."**

Ally McBeal, First Season, Episode-23 "These are the Days"

1.1 Background

Recently, information technology has been studied in a lot of research areas, and many products are undoubtedly accepted all over the world. Other technologies (e.g., mechanical engineering, electric engineering, electronic engineering, chemical science) also have gave us rich and comfortable lives. Current researchers must be aware of general needs and a right direction what technology really should be achieved in the future. Hence, I first start with a discussion what future information technology would become an essential thing for our everyday life in this section.

Miniaturization of computers, increase of storage size of the computers, and networked computers would give us key technologies being representative of the 21st century. These progressive technologies would make our everyday life advanced information society to enrich their individual life and collaborative life in the real world. A small and light-weighted computer can be worn by a user. Mann, S. (1997), for example, proposed a user who wears a CCD camera and sensors to record his/her experiences in his/her everyday life. This system allows the user to get information he/she needs anytime and anywhere. This type of human-centered computer technology is called "wearable computing." Wearable computing technology must be aware of both the user's internal (desire, emotion, health, action, etc.) and external (goings-on, temperature, other people, etc.) state anytime. Also, Kidode, M. (2002) has developed an advanced media technology project named a Wearable Information Playing Station (WIPS). My study on augmented memory in this thesis is a part of the WIPS project. The contributions of the WIPS project would be some of the key technologies for wearable computing in our everyday life. Also, real world-oriented appliances spread in the real world and linked with the Internet enable people to use them anytime and anywhere seamlessly. Weiser, M. (1991) proposed this type of distributed and invisible computing concept called "ubiquitous computing." The Aware Home Research Initiative is also directly inspired by the same concept (Kidd, C.D., et al., 1999; Abowd, G.D. & Mynatt, E.D., 2002; Tran, Q.T. & Mynatt, E.D., 2003).

In order to achieve the enriched our everyday life, a support of human's memory activities using the wearable computing and/or ubiquitous computing would be one of essential issues. Human memory does have fatal errors for our comfortable everyday life in essence. First of all, we shoud know what errors are in human memory activity. Scientific psychological analysis for memory-aid has been studied extensively in recent years. Psychological results show that the human brain can cause mistakes in either of encoding, storing, or retrieval process. Nickerson, R.S. & Adams, M.J. (1979), for example, showed that most American test subjects could not draw even 50% of the design of a 1-cent coin when they were asked to remember the coin visually. Brewer, W.F. & Treyens, J.C. (1981) investigated what strategies people can use to memorize environmental contexts in a room. The prospective memory research area gives us knowledge for memory-aid. Ceci, S.J. & Bronfenbrenner, U. (1985) analyzed the role of monitoring. By writing down the content of an intention when the intention rise in your conciousness, the moery-aid (monitoring) minimizes the influence of a retroactive inhibition (Intons-Peterson, M.J., 1996). The meta cognition using the memory-aid, however, has a limitation for automatic remembering of the intention (Wilkins, A.J. & Baddeley, A.D., 1978; Levy, R.L. & Loftus, G.R., 1984). In the remembering of intentions, the information we get from cues is important for us to recall the intention (Einstein, G.O., at al., 1998). McDaniel, M.A., et al. (1998) showed that the processing of the prospective memory is the concept-driven processing, and is not the perception-driven prosessing. West, R. & Craik, F.I.M. (1999) investigated the role of cue accessibility and cue sensitivity, and showed the effectiveness of the cue sensitivity. *Mäntylä*, T. & Sgaramella, T. (1997) reviewed the "Zeigarnik effect," and investigated that the encoding process seriously influences the latter recall that happens after the process. In contrast, the depth of the encoding decreases the performance of prospective memory (Schaefer, et al., 1998). In the prospective memory, "intention-superiority effect" is shown as one of good memory-aids that relatively gives high representation acitivities of intentions during a storing period (Goschke, T. & Kuhl, J., 1993; Goschke, T. & Kuhl, J., 1996; Marsh, R.L., et al., 1998; Marsh, R.L., et al., 1999). There are applicable contributions in the prospective researches, e.g, air traffic control (Vortac, O.U., et al., 1993; Vortac, O.U., et al., 1995; Vortac, O.U., at al., 1996) and space mission (Sauer, J., et al., 1999; Sauer, J., 2000). In the reserch area of human error (Reason, J., 1990), human errors are composed of "slip," "lapse," and "mistake." In terms of "slip," more exactly "action slips," the representative investivation is given by Reason, J. (1979). Models for action slip are proposed (Norman, D.A., 1981; Reason, J., 1984; Heckhausen, H. & Beckman, J., 1990). Autobiographical memory relating to "self" is also one of important issues in psychology (Brewer, W.F., 1986; Robinson, J.A., 1986; Robinson, J.A. & Swanson, K.L., 1990; Nelson, K., 1993; Schacter, D.L., 1995; Cohen, G., 1996). Autobiographical memories are not equally destributed across a person's lifespan. This phenomenon is called "bump" (Rubin, D.C., et al., 1986; Jansari, A. & Parkin, A.J., 1996; Rubin, D.C., & Schulkind, M.D., 1997; Rubin, D.C., 2000). In the lifespan, Barsalou, L.W. (1988) pointed out flexible and multi-dimensional structure of the autobiograpchical memory, and Conway (Conway, M.A., 1992; Conway, M.A. & Rubin, D.C., 1993; Conway, M.A., 1996a; Conway, M.A., 1996b; Conway, M.A. & Pleydell-Pearce, C.W., 2000) also proposed a structural model of it. Additionally, researchers suppose that autobiographical memories have a narrative structure that are tied a cause event to an effect event (Bruner, J., 1987; Gergen, K.J. & Gergen, M.M., 1988; Robinson, J.A., 1992; Barclay, C.R., 1996; Brown, N.R. & Schopflocher, D. 1998a; Brown, N.R. & Schopflocher, D. 1998b). Autobiographical memory is important for us as "self defining memory" (Singer, J.A. & Salovey, P., 1993) and "life theme" (Csikszentmihalyi, M. & Beattie, O.V., 1979). In the autobiographical memory, researchers also discussed "inplicit theory" that past personal histories are reconstructed by current self consious for corresponding to the condition of the consious (Ross, M., et al., 1981; Ross, M., et al., 1983; Ross, M. & Conway, M., 1986; Ross, M., 1989; Ross, M. & Buehler, R., 1994). By the contributions, computer could augment the lack of autobiographical memories in his/her life. In order to improve reliability of eyewitness evidence, "cognitive interview technique (CI)" has been developed (Geiselman, R.E., et al., 1984). This technique is based on "encoding specificity principle" (Flexser, A. & Tulving, E., 1978) and "multicomponent view of memory trace" (Bower, G.H., 1967). Generally, iteration of recall facilitates us to recall events itself (Erdelyi, M.H. 1996). The iteration, however, gives us more false memories at the same time (Ceci, S.J., et al., 1994a; Ceci, S.J., et al., 1994b; Hyman, I.E.Jr., et al., 1995; Hyman, I.E.Jr. & Pentland, J., 1996; Hyman, I.E.Jr. & Billings, F.J., 1998;). This false memory problem basically relates to the memory support using a computer.

The technology of the computational augmentation of human memory aims at integrating computationally recorded multimedia data named, "augmented memory" (Rhodes, B., 1997), into human memory. The ultimate goal of augmented memory is to enable users to conduct themselves using these memories seamlessly anywhere, anytime in their everyday life. In particular, video data provides the user with strong stimuli to recall past events that they have experienced. My research is composed of several works used in implementing a video-based augmented memory. In the field of computational memory-aid, several representative works on wearable computing and ubiquitous computing exist. Jimminy, a Wearable Remembrance Agent, also supports human activities using just-in time information retrieval (Rhodes, B., 2003). Kawashima, T., et al. (2002) and Toda, M., et al. (2003) have developed an automatic video summarization system using visual pattern recognition methods for detecting user's location and action recognition with a view tracking device. The Mithril platform has also advanced over the years (DeVaul, R.W., et al., 2003). Lamming, M. & Flynn, M. (1994) have developed Forget-me-not, a prototype system as a portable episodic memory aid. This system records a user's action history using sensors implanted in a laboratory and active badges worn by users. The user can refer to his/her own history and easily replay a past event in a PDA.

The other augmented memory systems are described in this section. I consider each system capable of being one of the modules for an ideal augmented memory system. In the case of location-based memory supporting systems, Hoisko, J. (2000) developed a visual episodic memory prosthesis, which retrieves video data recorded at the place attached to certain IR tags. The Global Positioning System (GPS) and the Geographic Information System (GIS) are also used for a location-based memory-aid. Ueda, T., et al. (2002) proposed an automatic video summarization method using a wearer position detected by the GPS, with the user's body direction by showing his/her interests around various architectures. In the object-based augmented memory, Rekimoto, J., et al. (1998) proposed a system for browsing information that includes the physical context of a user's current situation using CyberCode (Rekimoto, J. & Ayatsuka, Y., 2000) tags. Their aim was to develop interactive technologies in physical environments. A proposed video replay system, DyPERS, stores a user's visual and auditory scenes (Jebara, T., et al., 1998; Schiele, B., et al., 1999). This system can retrieve a video clip using a signal that was explicitly registered by a user who pushes a button while he/she looks at an interesting scene. The VAM system by Farringdon, J. & Oni, V. (2000) detects a human face recorded previously, and displays information of the retrieved person. Kato, T., et al. (2002a; 2002b) proposed VizWear-Active, which includes a robust face registration method and a stable face-tracking method. Segmentation of information and summarization research has been studied because such research provides information necessary for wearers to be able refer to their recorded viewpoint video data. Aizawa, K., et al. (2001) proposed the system that summarizes a video scene from a head-worm camera by using brain waves. Healey, J. & Picard, R.W. (1998) developed StartleCam, which records video data triggered by skin conductivity from a startled response from a user. In the Affective Wearables, various types of sensors were investigated to analyze and detect expressions of emotion in a certain environment (Picard, R.W. & Healey, J., 1997). In addition, video segmentation methods in the following methods have been proposed: a statistical trajectory segmentation method (Gelgon, M. & Tilhou, K., 2002), and a wearer's attention to a moving/stable object (Nakamura, Y., et al., 2000). Augmented memory has been investigated from the point of view of important psychological perspectives by DeVaul, R.W., et al. (2003) and Czerwinski, M. & Horvitz, E. (2002). Context awareness (Clarkson, B. & Pentland, A., 2000; Clarkson, B. & Pentland, A., 2001) and the modeling of event structure, as in Ubi-UCAM (Jang, S. & Woo, W., 2003), are also important. Ueoka, R., et al. (2001) studied what kinds of sensors are necessary. Murakami, H. (2002) developed an editing system named the Memory Organizer that enables users to refer to memory with a multi-directional viewpoint of interests using weak information structures, as in CoMeMo (Maeda, H., et al., 1997). The ComicDiary (Sumi, Y., et al., 2002) automatically produces a comic style diary of the users' individual experiences.

1.2 Motivation

The motivation of this work is to realize computational augmentation of human memory in everyday life. Does people need memory augmentation? YES! Many people (elderly people, amnesiacs, my friends, my parents, other field researchers, company developers, and so on) wait for memory augmentation products as soon as possible. Actually, memory activities are the first triggered process for recollection of past events, current works, and future plans anytime in everyday life. However, errors tend to occur in memory activities. The errors in the memory activities make us give up a scheduled action. Supporting memory process, e.g., encoding, storage, and recall, has been important issue for a long time. Were any memory support tools provided for us? YES! Lots of memory support tools have already been produced, e.g., a personal day planner, a video/audio recorder, photo album, and diary. Recently, Personal Digital Assistants (PDAs) have been in widespread use. The aims of such kinds of products are to formalize a design and an implementation of augmented memory that enables users to employ human memory and augmented memory seamlessly anywhere and anytime in their everyday life.

- - - - - - - - - - - - - - *Ambiguous Concept and Definition of Augmented Memory*

Speaking as a user, the best relation between human and computer is what the user can regard as if the system of the computational human memory augmentation is a part of natural human memory system. What is computational augmentation of human memory? what is called "augmented memory"? Generally speaking, the answer is that the augmented memory was just another term of a memory-aid system. I do not think that the augmented memory is same as conventional memory-aid systems. I think that the user should seamlessly, naturally, and unconsciously access to human memory and augmented memory. What pattern of memory-aid is the true augmented memory? How do I make the concept of the augmented memory? Although the research area must answer these questions in order to achieve the ultimate goal of augmented memory. Neither the concept of augmented memory nor definitions of augmented memory have been clarified still enough. This thesis first discuss what is the augmented memory, and then redefine it.

- - - - - - - - - - - - - - - - - - - *Implementation Modules without Global Perspective*

Most developed systems have not been designed for system integration of functions of augmented memory. Any systems, which can support user's memory activity using only mono-function, are not the ideal augmented memory themselves. I definitely promise that any mono-functional memory-aid systems never become a killer application. Do you think that people use lots of memory-aid systems at the same time? How many computers, PDAs, and a cell phones must we have? In the memory activity, people tend to surf on the organic networks flexibly associated memory functions with themselves. Should we consider that a macro design of the idel augmented memory? Actually, the research field of augmented memory has not yet had any reliable macro designs. In order to achieve the ultimate goal of augmented memory, I guess that we first design a framework that can link elements of augmented memory to themselves.

In addition to the framework of augmented memory, it is also important to design elements of augmented memory. However, most conventional memory-aid systems are developed as memory-aid tools. Memory-aid tools means one of externalizing systems like a diary, a memo pad, and a personal day planner. Are conventional memory-aid tools truly computational augmentation of human memory? NO! We should study what kinds of conceptual design has a potential for seamlessly/invisibly augmenting user's memory, and how the design is translated into reality.

1.3 Overview of the Thesis

Now, the thesis is ready to introduce "**A Design and Implementation of a Framework for Augmented Memory**." The thesis is composed of four parts. The first part is "**Part I: Introduction**." The second part is "**Part II: The SARA Framework**." The third part is "**Part III: Implementation of SARA Modules**." The last part is "**Part IV: Conclusions**." Figure 1.1

illustrates the roadmap of the thesis. The map shows the relation among topics in the thesis.

- **Part II : The SARA Framework**

Chapter 3 first introduces computational augmentation of human memory (in common name, "augmented memory"). The second section explains the first written augmented memory by Rhodes, B. And then, the third section re-defines the augmented memory. The final section summarize the chapter.

Chapter 4 shows an overview of a framework of SARA (Sceneful Augmented Memory Album) for augmented memory. This chapter has two definitions to give a frame for design and implementation of augmented memory. One is categorized based on the type of fuctions for dealing with augmented memory. In terms of this kind of categories, the first section explains the following functional categories, **Memory Retrieval**, **Memory Transportation**, **Memory Exchange**, and **Memory Editing**. Another definition is associable categories for clarify a trigger of an augmented memory module and for facilitating a connect-ability among other modules. The second section introduces 5W1H-based structural categories, **Spatial** (Where), **Physical** (Who, What), **Temporal** (When), **Psychal** (Why), and **Behavioral** (How). Finally, summary is in the last section.

- **Part III : Implementation of SARA Modules**

This part introduces three kinds of SARA modules. The part is divided into three chapters, i.e., a module for a Spatical element, for a Physical Human element, and for a Physical Object element.

Chapter 5 introduces a location-triggered video retrieval module named **Residual Memory** as an example of associable categories of Spatial element. The module show one of functions for Memory Retrieval.

Chapter 6 explains a human-triggered video management module termed **Nice2CU** as a "physical human" operational category. The module means one of Memory Retrieval and Memory Editing functions.

Chapter 7 shows a object-triggered augmented memory management module, **Ubiquitous Memories**, as a "physical object" operational category. The module has one of Memory Retrieval, Memory Editing, and Memory Exchange functions.

- **Part IV : Conclusions**

This part first introduces contributions of this thesis. In the next section, I discuss further directions for realizing the SARA framework and for developing SARA modules. I give a concluding remarks to finish the thesis.

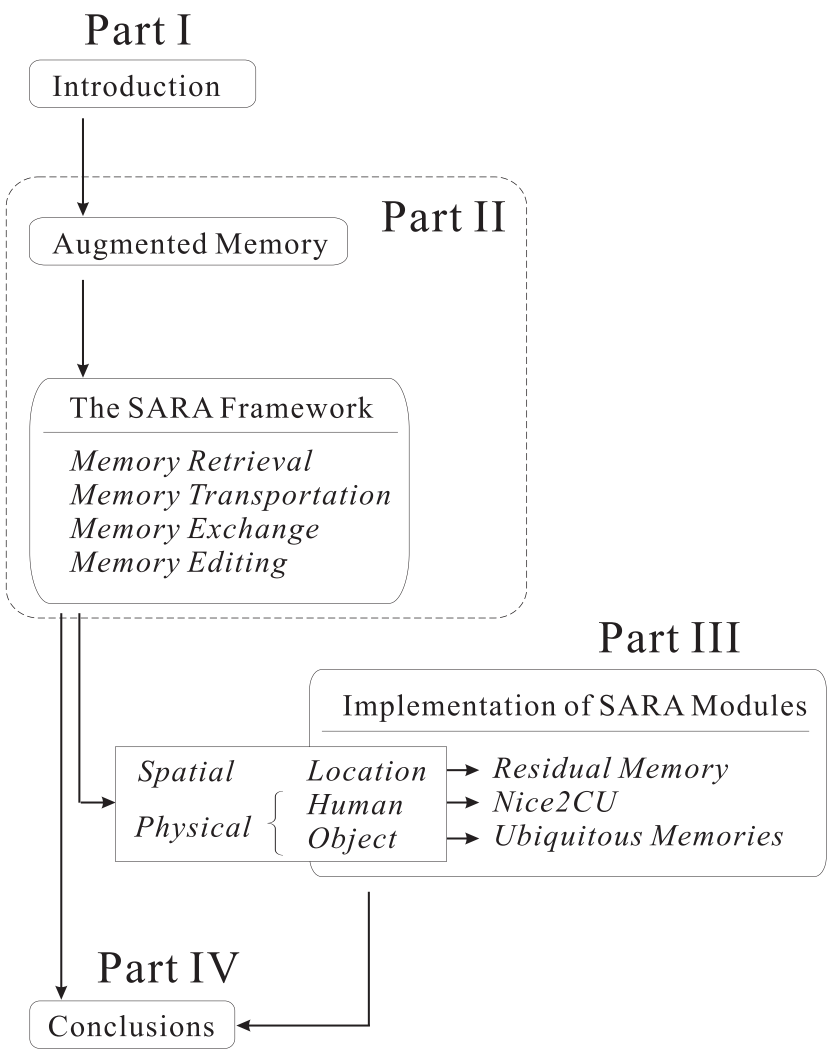


Figure 1.1. Roadmap of the Thesis

Part II

The SARA Framework

Chapter 2

Augmented Memory

In this chapter, I discuss what the "Augmented Memroy" is. The term of augmented memory is well known in a research community of wearable computing. The term is often used as a meaning of memory-aid systems on a wearable computer. Should we use the term as a common area to only limit to a domain for a proposing system? No, I do not think so. I strongly expect that the term of "'Augmented Memory" has a bigger potential than other terms meaning "memory-aid." In the first section, I introduce an original augmented memory proposed by Rhodes, B. Second, I discuss what components functions for features of the computational augmentation of human memory should have (the same as "Augmented Memory"). Third, I define a structure of augmented memory. I summarize this chapter in the last section.

2.1 Rhodes's Augmented Memory

Rhodes, B. has used the term of "Augmented Memory" as a system for computational augmentation of human memory. The term of augmented memory was first titled "The Wearalbe Remembrance Agent: A System for **Augmented Memory**" in the proceedings of the 1st International Symposium on Wearable Computers, October, 1997. Figure 2.1 is also written in a web page of the MIT wearable computing group. The figure is mainly described the Rebembrance *http://www.media.mit.edu/wearables/lizzy/augmented-memory.html*

Augmented Memory

One of the key differences between a wearable computer and the currently available palmtops is that wearables are always operational, tend to have sensors into their environement, and tend to have the ability to get information to the wearer even when the wearer doesn't expect it. This opens the door to a whole range of augmented memory applications specifically for wearable computers.

A simple example of augmented memory would be a traditional scheduling program which alerted you just before important meetings. Your wearable would simply whisper the information in your ear, or flash it to a heads-up display. More interesting are applications similar to the Rememberance Agent (RA) system being worked on by Rhodes, B., which as a user types or walks around continuously looks for documents with content relevant to the user's current situation. The file names or salient lines from these correlated documents are then continuously displayed at the bottom of the user's word processor in order of similarity. The system can find similarity based on the words currently being typed, people currently present, current location, and other physical information. On a wearable equipped with sensors like GPS, an indoor location system, face recognition, and speech recognition, the RA can give information based not only on what the wearer is typing but also based on his or her physical environment. For example, on meeting somone at a trade show or conference it could remind the wearer who this person was, what their vital information is, and bring up the notes taken at the last meeting with this person.

The RA can also be used to create a "group mind" by sharing databases of files. For example, the annotations and corrections one student makes over the course of a term may be very valuable to another student. The RA can index other people's notes, and automatically bring up their notes and expertise whenever it is most relevant to me. This can be especially useful for knowledge transfer and on-the-job training for new workers.

Last modified: Fri Nov 10 10:29:59 1995

Figure 2.1. Rhodes's Augmented Memory on the Webpage

Agent (RA) by Rhodes, B. The figure is composed of three paragraphs. The first paragraph describes the differences between a wearable computer and the conventional palmtops. The second paragraph shows a simple example of scheduling program and his RA. The last paragraph describes another ability of the RA for knowledge transfer creating "group mind." The figure also gives us information of equipment and techniques for augmentated memory. Wearables have sonsors that can recognize wearers environments, and they are implicitly supported their memory activities everytime and everywhere. It means that context aware technologies are important for accomplish a computational human memory augmentation. The figure also proposes an importance of technologies, e.g., locaton recognition, face regocgnition, speech recognition, and recognition of other physical information.

However, the figure does not describe what augmented memory is. Is the term of "Augmented Memory" only the role of a cool technical term for representing a wearable memory-aid system? I regret this situation of the term if the role of the term is so. We should embody the mechanism or concept of augmented memory if we viscerally use the term.

2.2 Definition of Augmented Memory

This section redefines the concept of augmnented memory. The aim of the section is to clarify preliminary ideal specifications of augmented memory for system developpers, users, and researchers. Suceeded conditions of augmented memory explained by Rhodes, B. are shown as follows:

Augmented memory runs...

- on a wearable computer and/or ubiquitous computers
- for supprting human memory activity in everyday life

Both are inevitable conditions for augmented memory. The former condition means that a platform that augmented memory enables a wearer to memorize and to recall an event anytime and anyywhere. The latter condition is a fundamental principle of augmented memory. After all, augmented memory functions must be immediately operated for the wearer anytime and anywhere even if he/she is unaware of a certain context included in an event he/she experiences. Additionally, I present ideal specification of augmented memory as follow:

Augmented memory gives the wearer...

- an recollect-able and appropriate information
- with low coginitive workload

The recollect-able and appropriate information does not have an automatic problem solver. Here, automatic problem solver means that augmented memory solves a certain problem without a person who wants to recall a certain memory with the problem. In the thesis, I assume that persons must have power of decision and action.

- *Definition*

Augmented Memory: *(1) "Augmented Memory" means augmenting human memory using a computing power. (2) The ultimate goal of the augmented memory is to trigger a user off his/her memory activity by "seamlessly" integrating his/her conventional human memory with the augmented memory in his/her "everyday life." (3) The augmented memory directly activates his/her memory activity activities (memorization/recollection) using a computing power, or indirectly supports his/her cognitive trait for the activities. (4) The augmented memory gives the user with lower cognitive workload of memory activities. (5) The augmented memory has a set of information, which triggers his/her memory activity, and functions that generates the information by measuring his/her inner status, activities, and surroundings with wearable sensors and/or ubiquitous sensors. The information is termed "Augmented Memory Source (AMS)." The function is named "Augmented Memory Function (AMF)." The augmented memory also has modules to operate the functions for triggering his/her memory activity. (6) The augmented memory employs anything as the storage of the information and the functions. (7) The augmented memory enables users to exchange their own augmented memories. (8) The augmented memory gives him/her means for editing his/her augmented memory. (9) The augmented memory enables the user to associate activated information by a certain function with a certain information to activate information appropriate for the next step of the association.*

- - - - - - - - - - - *Detailed Explanation of The Definition of Augmented Memory*

The augmented memory has been defined on the above paragraph. We however have still some difficulties to understand the definition because of a few explanations. Here, this section respectively explains the definitions in detail.

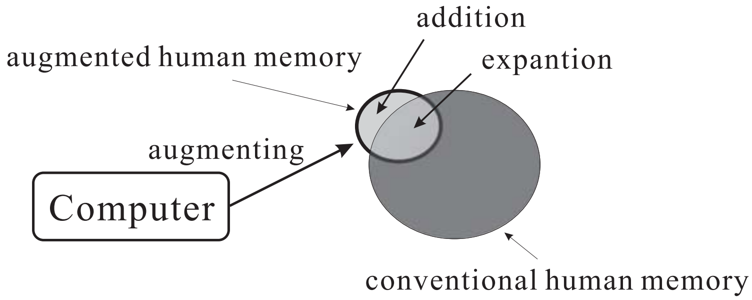


Figure 2.2. Augmenting Human Memory with a Computer(s)

Figure 2.2 illustrates the definition (1). Conventional human memory is the base of memory augmentation. A computer supports human memory by analysing his/her activity and envrionment. "Augmenting" has two meaning for supporting human memory activities. One is that the computer directly augments natural human memory. In other words, it means a mechanism "expansion" of natural human memory activities. Another is a virtual "addition" of new feature human beings never acquire in the evolution process. The "expansion" and "addition" are totally called "augmented (human) memory."

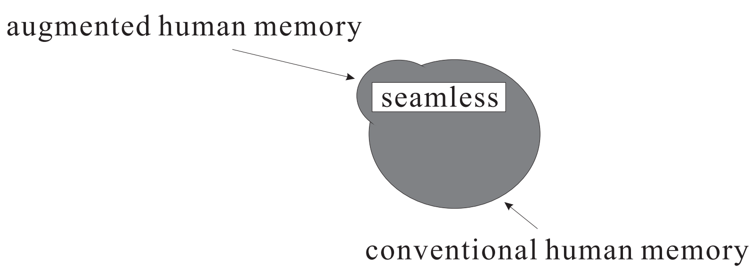


Figure 2.3. Ultimate Goal of Augmented Memory

Figure 2.3 shows the definition (2). Ultimate goal of computer-human interface is to be transparent interface to its user. Transparent interface means that the user is not regard to be aware of the interface at all. In terms of augmented memory, there is the similar ultimate goal. The goal is to integrate seamlessly the user's memory activity in the conventional human memory with the augmented human memory.

Figure 2.4 depicts the definition (3). There are two activating strategies in

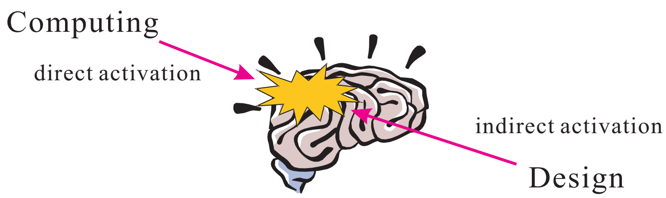


Figure 2.4. Direct Activation vs. Indirect Activation

the definition. One is a direct activation of memory activirty by the output of the computer. For example, the computer shows audio and/or visual information enabling the user to drive his/her memory activities. Another is an indirect activation of memory activity using an induction design of the memory augmentation. The system design could give a strong effect for a memory activity although this kind of system design limits the number of user's activities.

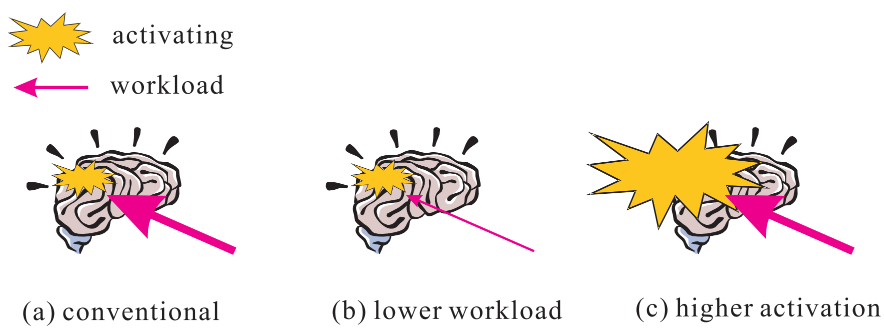


Figure 2.5. High Performance Cognitive Workload

Figure 2.5 shows the definition (4). Inefficient cognitive workload of memory activity prevents a progress of augmented memory from realizing the ultimate goal. Effective memory activity is accomplished by the following two ways. Figure 2.5(a) illustrates a conventional (base) model of memory activity. Figure 2.5(b) shows a lower workload model. In comparison with the base model, the lower workload model has the same strength of activation using an external stimuli for augmenting memory. Also, Figure 2.5(c) depicts a higher activation model. The model motivates higher memory activity than the base model although the model shows the same workload of the base model.

Figure 2.6 shows the definition (5). Augmented memory is composed of Aug-

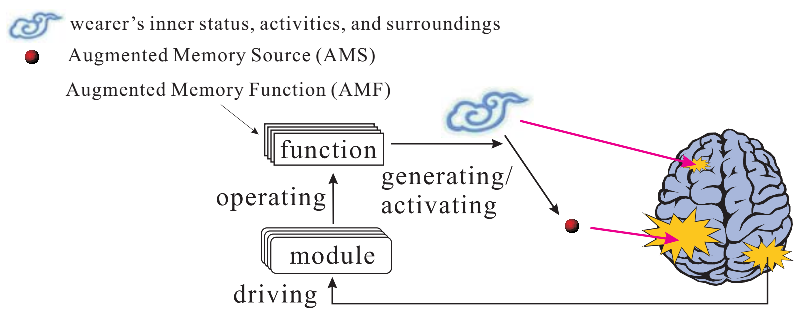


Figure 2.6. A set of Information, Functions, and Modules

mented Memory Sources (AMSs), Augmented Memory Functions (AMFs) and modules. An AMS is an element for driving the user's memory activity. The AMS generates stimuli for memorization and recollection greater than raw information captured by wearable sensors and/or ubiquitous sensors. A certain AMS is generated by a certain AMF that has a purpose of a domain specific trigger for memorizing or recollecting of the user's experiences. The module is drived by the memory activity or surrounding. The module operates AMFs for a designed memory support strategy.

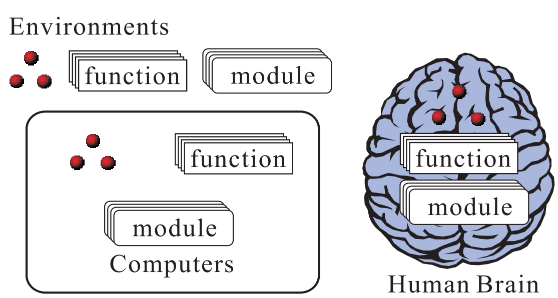


Figure 2.7. A storage of Information, Functions, and Modules

Figure 2.7 illustrates the definition (6). The augmented memory employs anything as the storage of AMSs, AMFs, and modules. It means the augmented memory can employ environments, human brains, and computers as a storage of AMSs, AMFs, and modules. AMSs, AMFs, and modules in the environment and human brain are virtually designed. The modules, however, run on only the wearable and/or ubiquitous computing devices because they are implemented by augmented memory module developers. The definition totally say that you can only implement AMFs and modules, and reorganize AMSs in the computers directly. AMSs, AMFs, and modules in the environment and the human brain can be virtually and indirectly controlled by AMSs and AMFs in the computers.

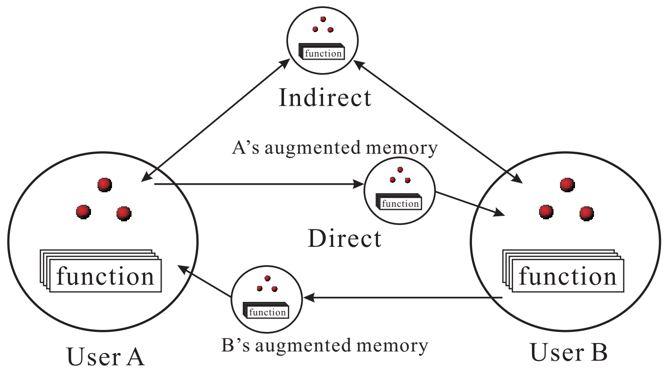


Figure 2.8. Exchanging of Augmented Memories

Figure 2.8 shows the definition (7). A user can give his/her own AMSs and/or AMFs to other users. These AMSs and AMFs are not only enable the target user to memorize/recall the owner's experiences for interpreting a certain experience from other view points, but also the function of the target's augmented memory is newly expanded by the exchanging AMFs. The user can select two augmented memory exchange methods. One is a direct method. This method helps a privacy-based on-demand exchange of his/her augmented memory. Another is an indirect method. The aim of this method is to open a public-based share of users augmented memories.

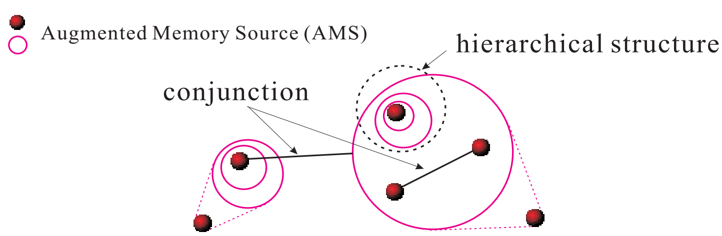


Figure 2.9. Editing Augmented Memories

Figure 2.9 depicts the definition (8). "Editing" in the augmented memory is to clarify a relation among AMSs. An AMF for editing links an AMS with another AMS. Here, the most important thing to explain the definition is that AMS has self contained structures. AMFs generate any kinds of self contained AMS and can operate any levels of AMS as the same types of AMS. This structural AMS including liked AMSs makes higher memory activity of a user than either AMS of the linked AMSs.

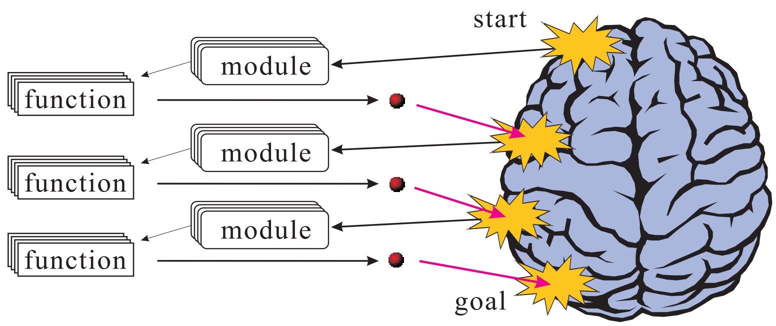


Figure 2.10. Association Process of Augmented Memory

Figure 2.10 shows the definition (9). People tend to forget what would be a right trigger for directly motivating an acquired memory activity. You, however, can pick up the other strategy to get the memory activity. That is "Association," a user memorizes or recollects an experience by pursuing a chain of associated experiences from the beginning trigger.

- *The Augmented Memory System*

Figure 2.11 totally shows the representation of the augmented memory system from the viewpoint of the idea world. There are the user A's and B's world. The world is composed of three domains: the human brain, the environment, and the computer. AMSs are respectively linked with other AMSs. Memory activities (memorization/recalls) are activated everywhere. (I) A certain module observes an certain memory activity. (II) The module then operates a function appropriate for the next memory activity. (III) The selected function drives a AMS. (IV) The AMS finally activates the right memory activity for him/her. (V) If he/she explicitly drives another module, this operation is called "association." The user (A/B) can exchange his/her own AMSs and AMFs with another user (B/A).

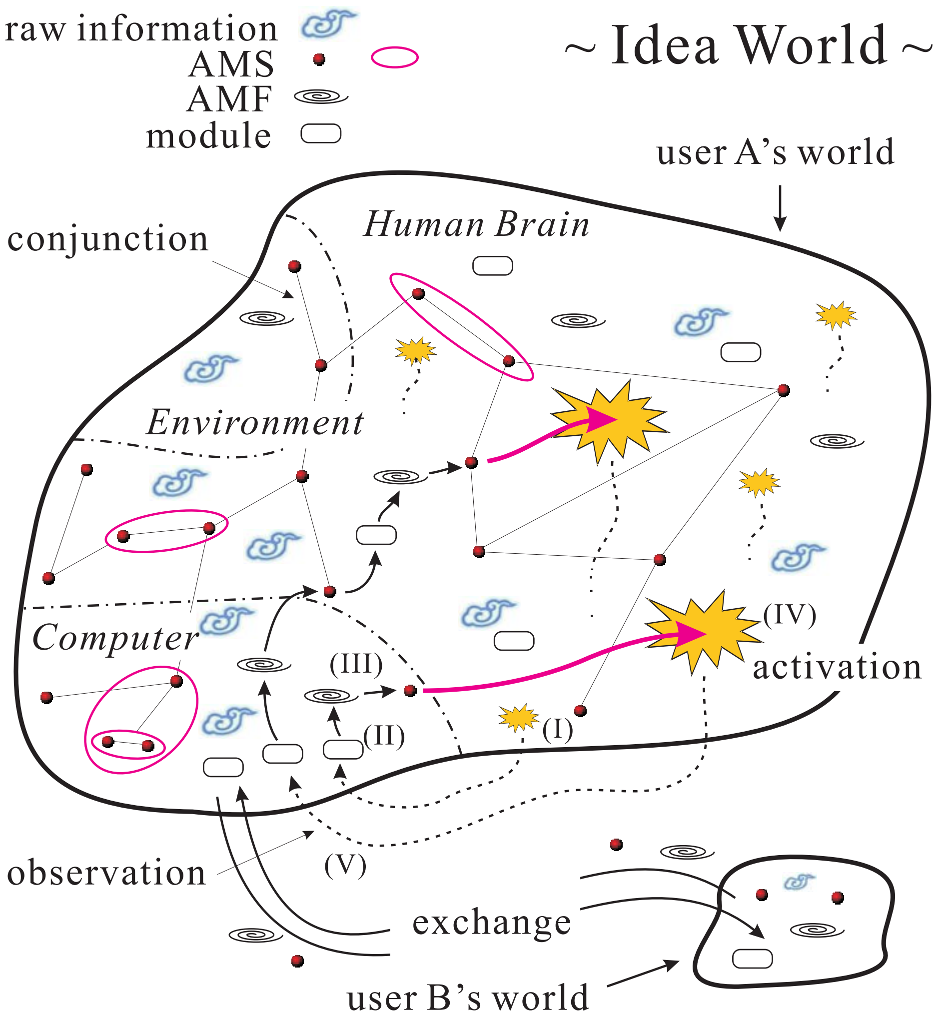


Figure 2.11. Idea World of The Augmented Memory System

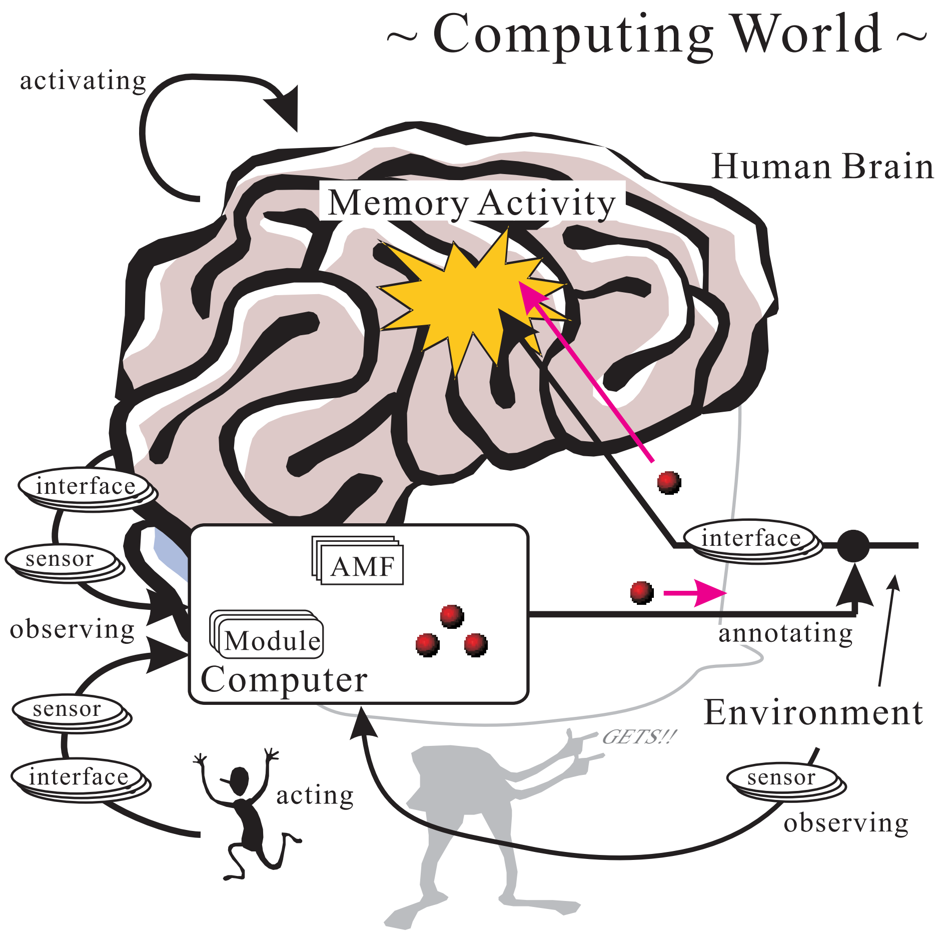


Figure 2.12. The Augmented Memory System

On the other hand, Figure 2.12 shows how the augmented memory system is constructed from the viewpoint of computing side. A computer for memory augmentation is composed of AMSs, AMFs, and modules. The computer also detects environmental information and human's activity from biometric sensors, motion sensors and user operations. Augmented memory modules choose appropriate AMFs for managing AMSs. AMSs are annotated to the environmental information, and then the user is driven his/her memory activity of a required event. The activity repeatedly motivates new activities by pursing associations. His/her reactions to the activity also make state transition of environment. After all, the transition triggers a new motivation of a behavior with a memory activity.

2.3 Ideal Specifications

Memory process from input through output is basically divided into three processes; "Encoding" process, "Storage" process, and "Retrieval" process. In order to accomplish augmented memory, the following issues lie on each process.

- Encoding 1: Automatic/Semi-automatic/

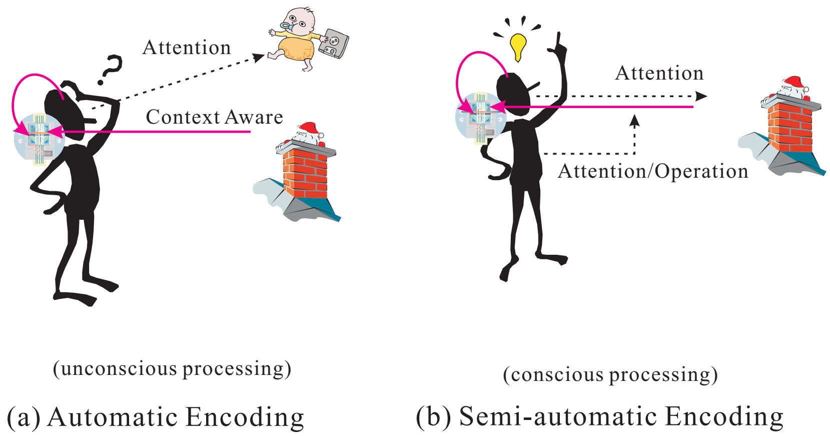
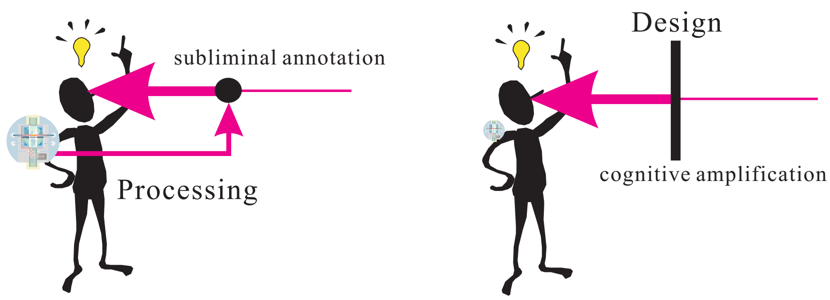


Figure 2.13. Automatic and Semi-automatic Encoding of Augmented Memory

Figure 2.13 shows "Automatic" and "Semi-automatic" encoding process. In the encoding process, "Automatic" is ideally the best method to encode acquired information if augmented memory can do. "Automatic" should collect not only information awared by the person but also unawared information at the moment of encoding. However, the method cannot employ his/her conscious intention. Hence, "Semi-automatic" is also an important function to resolve the problem.

- Encoding 2: Subliminal/Cognitive/

Figure 2.14 shows a model of "Subliminal" and "Cognitive" encoding process. "Subliminal" in the encoding process means that a person's brain is stimulus for memorizing an event under subconscious level when the event is encoded as a computational augmented memory. The subliminal encoding process is a proactive strategy for alleviating cognitive (recollection) workloads in a retrieval process.



(a) Subliminal Encoding (b) Cognitive Encoding

Figure 2.14. Subliminal vs. Cognitive Encoding

In contrast, "Cognitive" encoding process is the memorization strategy using human's cognitive characteristics. The advantage of "Cognitive" encoding prosess is that augmented memory does not require any computational techniques. However, the function must have an ability of cognitive induction by an ideal encoding design.

- Storage: Personalize/Communitize[1] /

Figure 2.15 shows "Personalize" and "Communitize" process. All persons perfectly experience an individual event even if they are at the same place, and see the same scene. "Personalize" is the most important things in the storage processes. Personally optimized memory is useful for the person who experienced an optimized event. Personalized memory is then commutized for communities to which he/she belongs. After all, optimizations on all layers are for having memory in common among each person's life (not for resolving). Persons can make good use of organized memory, and then they provide new memories for organizing again to us. I have named this cycle "Experience Recycling."

Figure 2.15 also shows an optimizing cycle for a person. The aim of organizing memories is to finally provide useful personal optimum memories to all person each. Augmented memory aiming for "Experience Recycling" would need not

[1] Here, "Communitize" is a coined word in the thesis. "Communitize" is composed of "Community" + "tize" to verbify for a meaning of "organize a community."

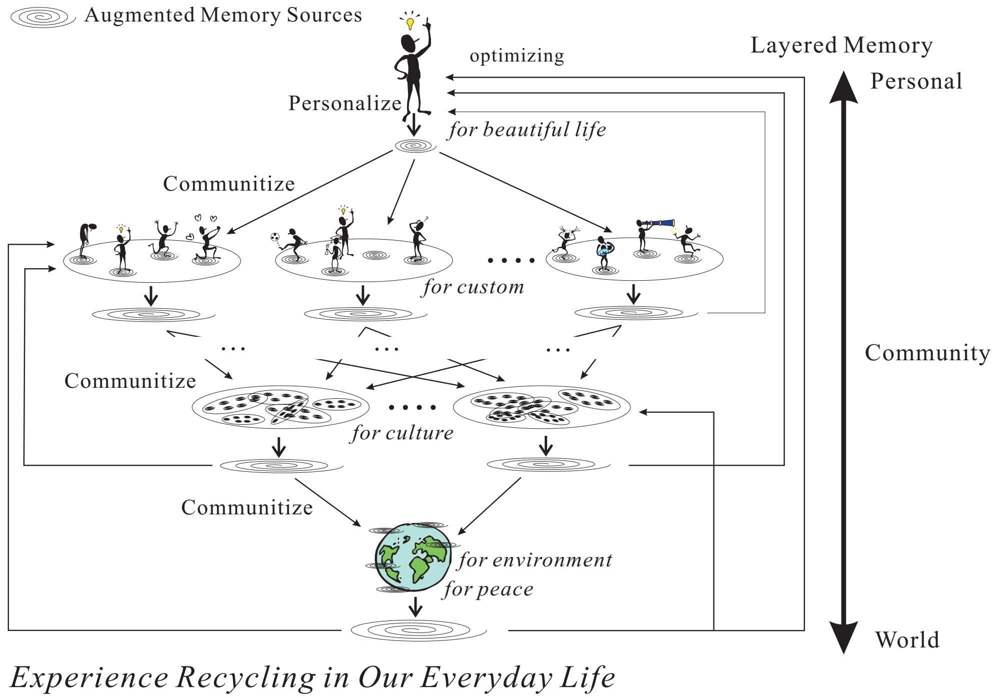


Figure 2.15. Organizing Augmented Memories for Experience Recycling

only organizing mechanisms but also personally optimizing mechanisms adapting person, requisition, and situation.

- Retrieval: Subliminal/Intuitive/

Figure 2.16 shows a model of "Subliminal" and "Intuitive" retrieval process. In the "Subliminal" retrieval process, a person recalls under subconscious level. However, automatic "Subliminal" retrieval process tends to lead on a wrong recollection of the person, and then he/she would act with an incorrect choice. Moreover, I have doubts that the person can get certainty with the "Subliminal" retrieval. In contrast, "Intuitive" retrieval process is based on a person's awareness of the support. The person can judge retrieved information with the "Intuitive" process. However, automatic "Intuitive" retrieval process would cause person's overload without the right timing.

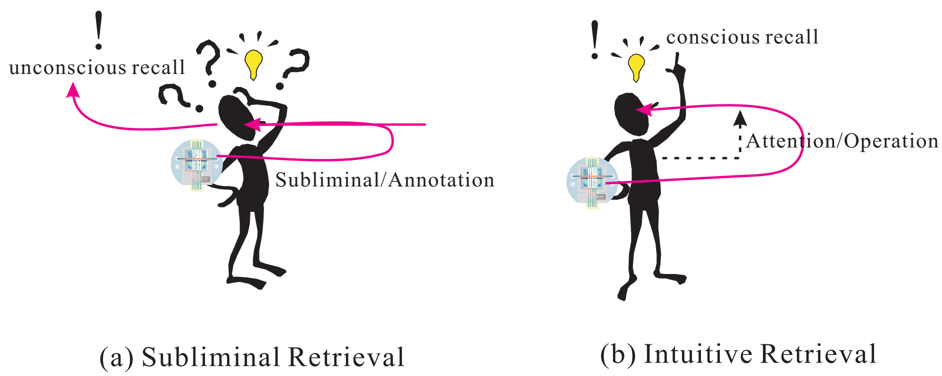


Figure 2.16. Subliminal vs. Intuitive Retrieval

2.4 Summary

In this chapter, I first introduced Rhodes's augmented memory. However, Rhodes's introduction has not been completed the explanation of augmented memory, and it's so abstract. Also, I proposed a new definition of augmented memory.

However, two problems are remained: 1) How can users operate augmented memory? 2) What kinds of functions does augmented memory need to implement it? Next chapter introduces of a framework named SARA (Sceneful Augmented Remembrance Album).

Chapter 3

SARA: Sceneful Augmented Remembrance Album

In this chapter I introduce a design frame for a computational memory-aid useful in everyday life. The aim of this chapter is to develop the design frame for a video-based albuming system for augmenting human memory. I propose a framework for an augmented memory albuming system named *SARA*[1] (Sceneful Augmented Remembrance Album), used to realize computational memory-aid in everyday life (Kawamura, T., et al., 2002b; Kono, Y., et al., 2003):

- A user's viewpoint images are always observed.
- The observed images along with the data observed by other wearable sensors are analyzed to detect contexts.

[1] *SARA* is named from the following sentenses of "the tale of the Heike." Note that "*śā*la" shows the same pronunciation of *SARA* in Japanese Hepburn system katakana. (in Japanese)「祇園精舎の鐘の声 諸行無常の響きあり 沙羅双樹の花の色 盛者必衰の理をあらはす おごれる人も久しからず 唯春の夜の夢のごとし たけき者も遂にはほろびぬ 偏に風の前の塵に同じ」『平家物語』/ (in English) "The sound of the Gion Shōja bells echoes the impermanence of all things; the color of the *śā*la flowers reveals the truth that the prosperous must decline. The proud do not endure, they are like a dream on a spring night; the mighty fall at last, they are as dust before the wind." (『*The Tale of the Heike*』: McCullougb, H.C., 1988)

- The observed images are stored with the detected contexts as the user's augmented memories.
- The stored data can be additionally annotated/indexed by the user for later retrieval
- The user can recall his/her experiences by reviewing a stored video, which is retrieved by consulting an index, via a Head-Mounted Display (HMD), i.e., both annotating and indexing data automatically stored and manually annotated information.

3.1 Operational Category

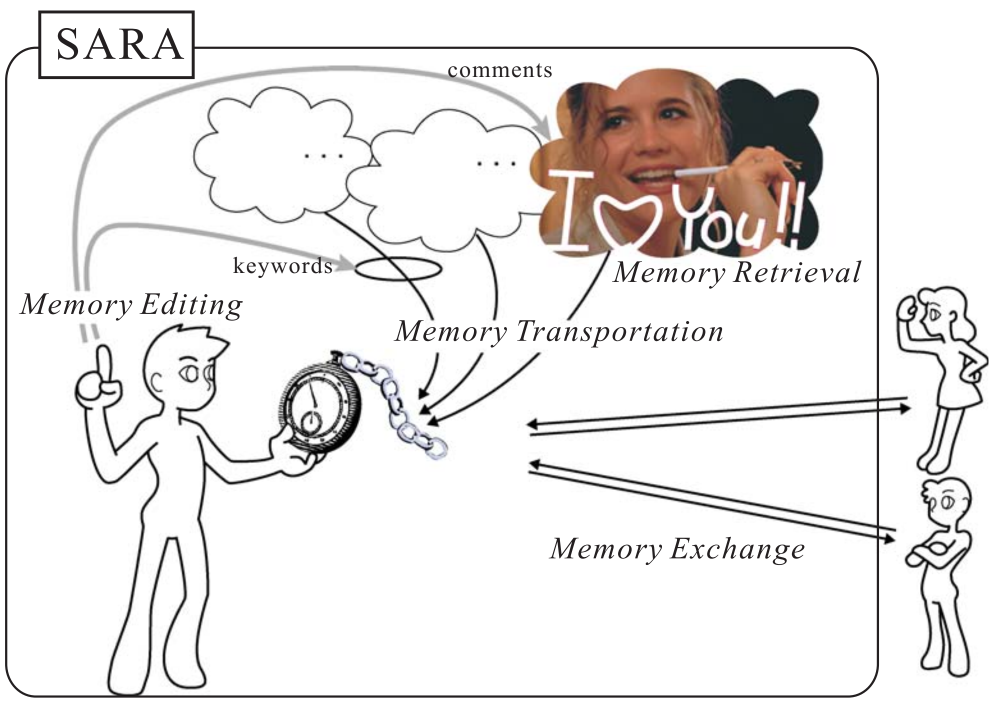


Figure 3.1. Basic Function for Operating Augmented Memories on SARA

Figure 3.1 illustrates the overall architecture of the proposed augmented memory albuming system. I believe that the following four functions are essential to realize *SARA*:

- **Memory Retrieval**

A memory retrieval function provides a user with the ability to retrieve appropriate multimedia data from a huge multimedia database termed "augmented memory source." A human chooses his/her best remembrance strategy from various remembrance strategies and applies this strategy to the context. I design and implement the following three modules for the Memory Retrieval: the "*Residual Memory*" for supporting a location-triggered user's recollection and the "*Nice2CU*" for a person-triggered one. "*Ubiquitous Memories*," mentioned in Memory Transportation, also works as an object-triggered memory retrieval module. In Chapter 4, I introduce the "*Residual Memory*" module. Also, I explain the "*Nice2CU*" module in Chapter 5. Lastly, I design and implement the "*Ubiquitous Memories*." in Chapter 6.

- **Memory Transportation**

A memory transportation function provides a user with the ability to associate either the event he/she is enacting or the augmented memory retrieved by one of the memory retrieval modules, with other features. By using such a function, the user is able to rearrange his/her memories for later retrieval.

- **Memory Exchange**

A user can augment his/her problem-solving ability by referring to others' experience if this experience is properly associated with the given problem. Additionally, the "*Ubiquitous Memories*" helps its users exchange their experience. By accessing a real world object, a user can view all the augmented memory sources associated with the object if the owner of each memory has approved other users viewing the object. I explain the "*Ubiquitous Memories*" module in Chapter 6.

- **Memory Editing**

A person remembers an event he/she has experienced with conceptual/linguistic indexes. A memory editing function provides a user with the ability to make annotations to his/her memories by adding keywords or free-hand (for instance, this means "*I♡You*" in Figure 3.1) comments, or by reordering his/her memories based on his/her own intentions. Our research group has developed a vision interface, which is named a Wearable Virtual Tablet (Ukita, N., et al., 2002), using only the user's fingertips The "*Nice2CU*" enables the user to edit his/her personal profiles for a meeting and a conversation with others by updating his/her own personal profiles. Using the "*Ubiquitous Memories*," the user can link augmented memories with real world objects. He/she can also hold and convey the memories to the linked objects. I explain the "*Nice2CU*" module in Chapter 5. I explain design and implementation of the "*Ubiquitous Memories*." in Chapter 6.

3.2 Associable Category

Associable categories gives wearer awareness what kinds of trriger he/she can use on a certain module. The categories also provide a frame to associate each modules using intuitive representations for users and developers. Figure 3.2 is an expanded structure elements of the world basing on the 5W1H (Who, What, When, Where, Why, How) model (Kawamura, T., et al., 2001a; Kawamura, T., et al., 2002a). Each element of the 5W1H model respectively corresponds to Physical Human (Who), Physical Object (What), Temporal (When), Spatial (Where), Psychal (Why), and Behavioral (How). The model is composed of a static world and a dynamic world. The static world is a sliced (picked up) dynamic world.

- *Static Base Elements*

- **Spatial (Where)**

"Spatial" element is the most important element in the structure model. Spatial element can exist without the time concept. Crucial example is shown as

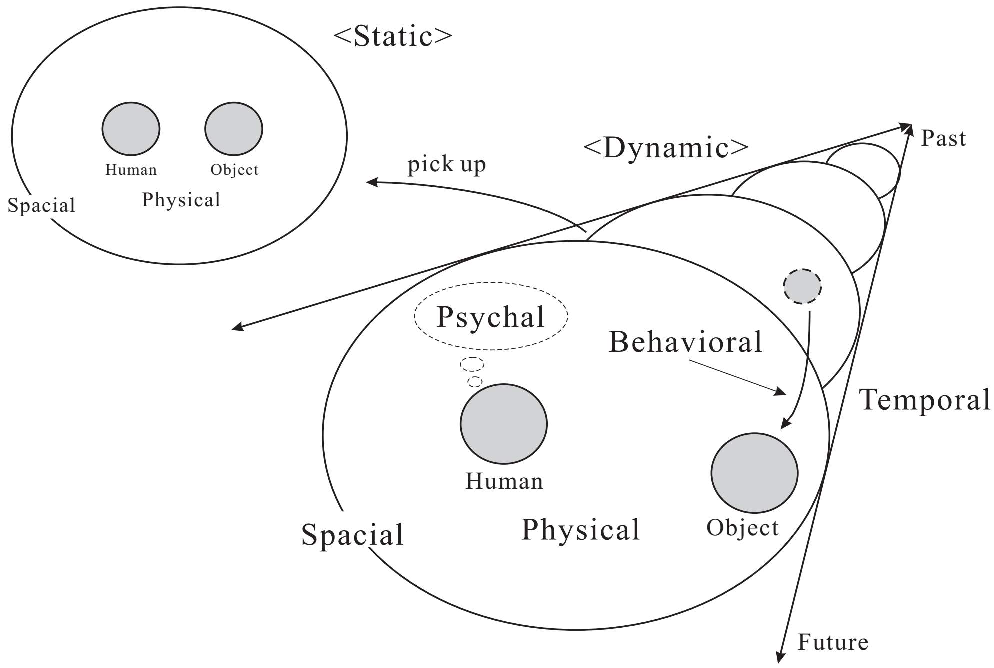


Figure 3.2. A Structure for Categorizing SARA Modules

"What was going on here?" This element can exist both a static world and a dynamic world. Most of contexts in the memory activities are tied to the Spatial element. Contexts specialized in the information space are none of the Spatial element's business. Actual spatial information exists in the real world. Real world that we live gives us stimuli for recollection everywhere. Everywhere stimuli facilitate intuitive information retrieval. Spatial information is the best trigger for recalling past events because people can easily recall the information at first This thesis introduce the "Spatial" element in Chapter 4.

- **Physical (What, Who)**

"Physical" element is embodiment of the real world because of configurational relations among physical elements. This element also exist both a static world and a dynamic world. Additionally, physical element is composed of "Physical Object" and "Physical Human." The difference between the physical object and the physical human is that only the physical human has an after-mentioned "Psychal" element. In a narrow sence, physical human contains all creature.

Physical objects make space visual speciality in the real world. Physical humans have human relations. Chapter 4 and Chater 5 respectively design and implement the augmented memory modules on the “Physical” element.

- *Dynamic Additional Elements*

“Temporal” element and “Behavioral” element are in the dynamic space at the same time. Time acts as host of transition, but time is not a factor of transition itself. In contrast, behavior gives context transitions in the real world using physical elements. However, behavior does not have a meaning of relative relations. Hence, “Temporal” element is inseparable from “Behavioral” element.

- **Temporal (When)**

Generally speaking, “Temporal” element gives dynamic events. However, an important discussion point of “Temporal” element is that time does not unilaterally generate any activities in the real world. Or, time is made by relative situation transitions activated by static base elements. This idea does not have any powers to reduce the existence value of “Temporal” element. Temporal element presents macro relative patterns of “Behavior” element. Time element shows the following kinds of patterns:

- **Tense:** Past, Present, Future

- **Behavior:** Cycle, Rhythm, Tempo

- **Elapse:** Gain, Hark back

- **Period(Tense):** Perfect, Progress, Stagnancy

Temporal element also gives us cause-and-effect relationship in the real world. Cause-and-effect relationship is one of time sequential information. This duality between cause and effect is one of the most important points of “Temporal” element. Relative relationship is also one of them.

- **Behavioral (How)**

In contrast to "Temporal" element, "Behavioral" element explains micro state transition itself. "Behavioral" element includes not only physical action but also other state transitions occurred by "Temporal" element. Here, in order to represent "Behavioral" information, we must decide a variation axis, a starting point and an ending point. For instance, temperature (a variation axis) in a room is increasing (start point: low, end point: high). The start point and the end point give us a direction of action information. In the thesis, I employ a continuous state as one of behavioral information. The continuous state can be measured using a comparison between the target information and other states that have been changing.

- **Psychal (Why)**

"Psychal" element is the most special element in the associative model. Psychal elements only exist in the brain of physical human. The concept is allowed if the element virtually associate with the physical human. Psychal information would be measured from physical human body activities, or via memory externalizing tools. Especially, we could abstractly estimate from brain functional mapping. Psychal element is composed of emotion, intention, desire, and conscious.

3.3 Summary

This chapter shows two kinds of categories to construct the SARA framework for augmented memory. One is the operational categories for operating augmented memory. The categories give us somewhat narrow ranges of user interfaces on the concept level. However, the categories also show existence of other similar operational functions on Memory Retrieval, Memory Transportation, Memory Exchange, and Memory Editing at the same time. Another is the associable categories for clarifying trigger in an augmented memory module. Additionally, in terms of ultimate goal of augmented memory, the categories must be needed for "Association" among mono associable-functional modules at least.

In the next part, I introduce three kinds of implemented augmented memory module. The modules are fundamental modules on the code of the SARA framework. The modules also include all static base element, “Spatial” element, “Physical Object” and Physical Human” elements. This coverage of static base elements means that they would work as base for the ultimate goal of augmented memory.

Part III

Design and Implementation of the SARA Modules

Chapter 4

Residual Memory - Spatial -

Main issues of this chapter are to enable a user to employ an operation for a location-triggered video retrieval using a user's viewpoint image. In order to establish these issues, this chapter shows the following technical topics for computing a locational similarity between recorded video and the user's viewpoint image:

- A user's viewpoint image is always recorded by a camera that captures the image.
- The recorded video data is composed of location information and motion information.
- The recorded video data is segmented into scenes for a easy viewing to the user.
- Video scenes are retrieved with a high-speed video tracking method.
- The captured image is reduced a sensor noise for high-precision by computing an image similarity.

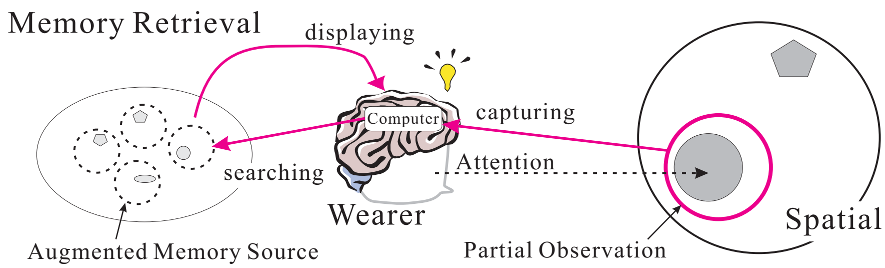


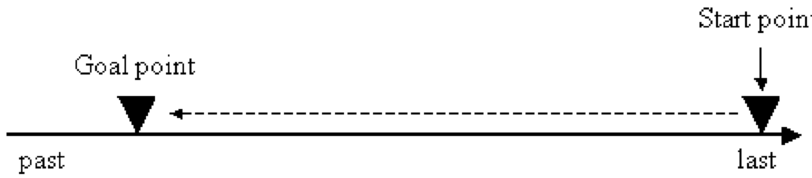
Figure 4.1. Locaion-triggered Augmented Memory Retrieval

4.1 Design

This chapter represents an example of the operational category of "Memory Retrieval" mentioned in Chapter 3. Also, this chapter shows an example of the associable category of "Spatial" in Chapter 3. Figure 4.1 illustrates a model of a location-triggered memory retrieval. Suppose that a user pays an attention to a certain space. The user tends to feel a certain deja vu phenomenon. A location-triggered augmented memory module captures the user's viewpoint image, searches the similar scene from augmented memory sources, and then the module shows the user the retrieved augmented memory source. The user augments his/her memory by recalling an event he/she feels the deja vu phenomenon.

I have developed a location-based video retrieval module named Residual Memory (Kawamura, T., et al., 2001; Kawamura, T., et al., 2002b). Residual Memory module assists a user in remembering an event that happened at the location. The Residual Memory provides associable video retrieval in a previous data set triggered by current video data that is captured from the wearable camera. The module needs high speed and an appropriate location-based video retrieval method for a user's on-demand request using his/her viewpoint image. Generally, the users head frequently moves, and moving objects also move in his/her view. The system must divide captured information into moving information and location information. In order to achieve these methods, the module has to track the user's head movements, detect moving objects in a scene, and remove these two motions from the video in the user's activities.

4.1.1 Conceptual Design



(a) From the latest image to goral image

(b) Previous tracking strategy

Figure 4.2. A sequential search and a tracking search

In the study, I assume that the recorded video data set is huge. If the system recorded a color video of the user's entire life, the size of all recorded color video images would be over 17 PB ($320(width) \times 240(height) \times 3(bytes) \times 30(frames) \times 60(seconds) \times 60(minutes) \times 24(hours) \times 365(days) \times 80(years)$). When the system links similar images and searches for a video image along linked video sequences, the retrieval time decreases. The similar image linking method, however, does not resolve the video data size problem. In Figure 4.2(a), if the system runs an image matching process sequentially from the latest video frame to the goal frame, then the system time would sometimes be over 1/30 of a second according to distance between the goal point and the latest video frame. A video retrieval system with a huge video data set, therefore, has to contain a function for resetting the starting point near the goal point, per interval, as shown in Figure 4.2(b).

Continuously recorded video data from the wearable camera has the following two characteristic features:

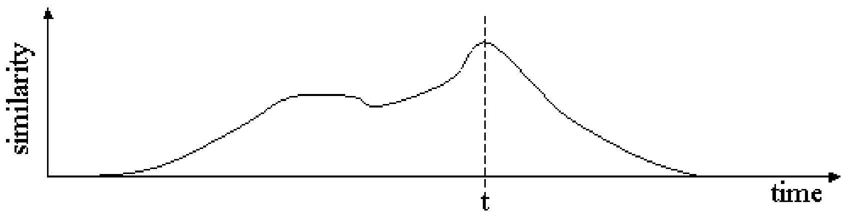


Figure 4.3. Indistint Video Scene Cluster

(1) Continuity: Adjacent images closely resembling each other. Similarity between the image of point t and another point changes smoothly with the expansion of a captured time difference (Figure 4.3). This feature creates indistinct video scene clusters.

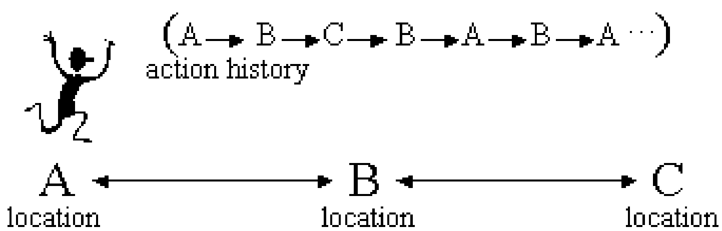


Figure 4.4. Spatial Video Feature

(2) Spatiality: Based on where the user is standing, similar images are captured by a similar spatial viewpoint. In any one day, for example, the user walks to places where he/she wants to go and. In Figure 4.4, for example, the user moves around places A, B, and C. A history of where the user walked can be shown as $A \rightarrow B \rightarrow C \rightarrow B \rightarrow A \rightarrow B \rightarrow A$. This example of action history represents how the user arrives at the same place within an odd interval.

In my video retrieval system, query images from the user's viewpoint are input with high frequency (the same frequency as the video input rate). By using high frequency input, the system can quickly track the changes of the user's request to refer a location-based associable video data. The user can always slightly change an image with his/her own body control and input a query image to the system at the same time. The system, however, must have the function of an on-the-fly adaptive video search. The system prepares itself using a background processing system for a user's unforseen choices of a scene in the huge video data set.

4.2 Implementation

4.2.1 System

- - - - - - - - - - - - - - - - - *Wearable Equipment for the Residual Memory system*

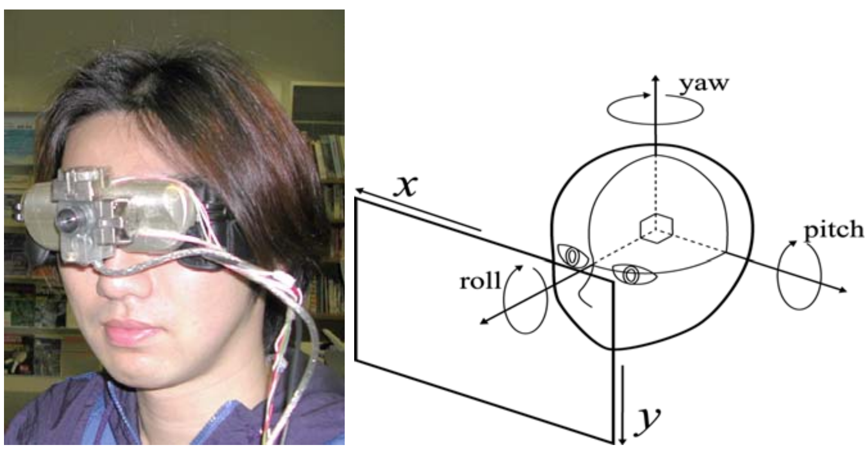


Figure 4.5. A Wearable Camera, Display, and Two Mono-axis Gyro Sensors

A user wears the OLYMPUS HMD, which is named EYE-TREK FMD-700. The HMD has a USB camera attached to the center of display. The camera is I-O DATA USB-CCD. The user's head motion prevents him/her from recognizing a moving object in a captured video from the wearable camera. I have employed two mono-axis gyro sensors (Silicon Sensing Systems), and placed these sensors in the center of the wearable camera as shown in Figure 4.5 (left). The two gyros can detect yaw-axis motion and pitch-axis motion (Figure 4.5 (right)).

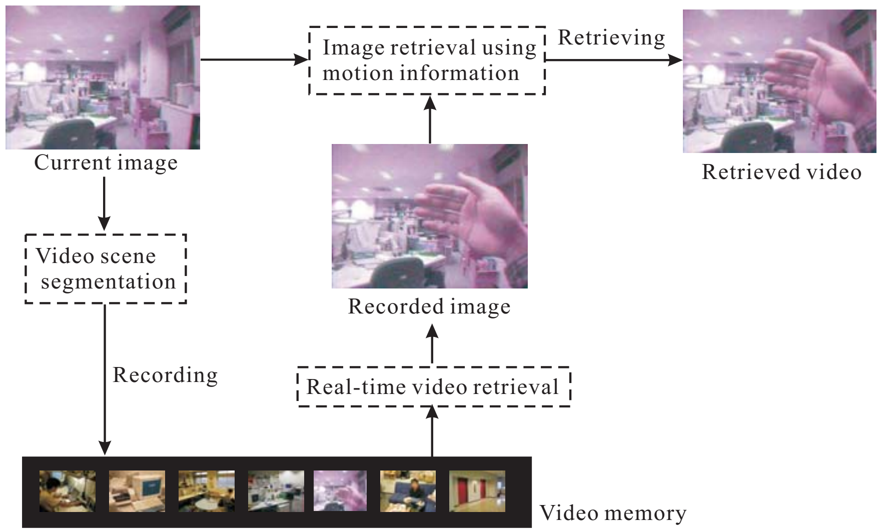


Figure 4.6. An Overview of Residual Memory Process Flow

The Residual Memory module operates the following three methods (Figure 4.6).

Motion information exclusion from video data is performed using the wearable camera by:

- tracking the yaw and pitch head movements with two mono-axis gyro sensors,
- tracking moving objects in a scene using a block matching method, and by
- excluding motion information by masking moving areas in the scene.

Video scene segmentation from continuous input video data is changed by:

- detecting scene changes continuously from current video data and two gyro data, and by
- indexing each scene for easy viewing.

Real-time video retrieval from large sequential video data is retrieved by:

- dividing small segments from video data for stable and high-speed video retrieval, and by
- retrieving an associable scene from a segment similar to the current video data.

In my video retrieval system, query images from the user's viewpoint are input with high frequency (the same frequency as the video input rate). By using a high frequency input, the system can quickly track the changes of the user's request to refer a location-based associable video data. The user can always slightly change an image with his/her own body control and input a query image to the system at the same time. The system, however, must have the function of an on-the-fly adaptive video search. The system prepares itself using a background processing system for a user's unforeseen choices of a scene in the huge video data set.

4.2.2 Image Matching Except for Motion Information

\- *Tracking Head Movements*

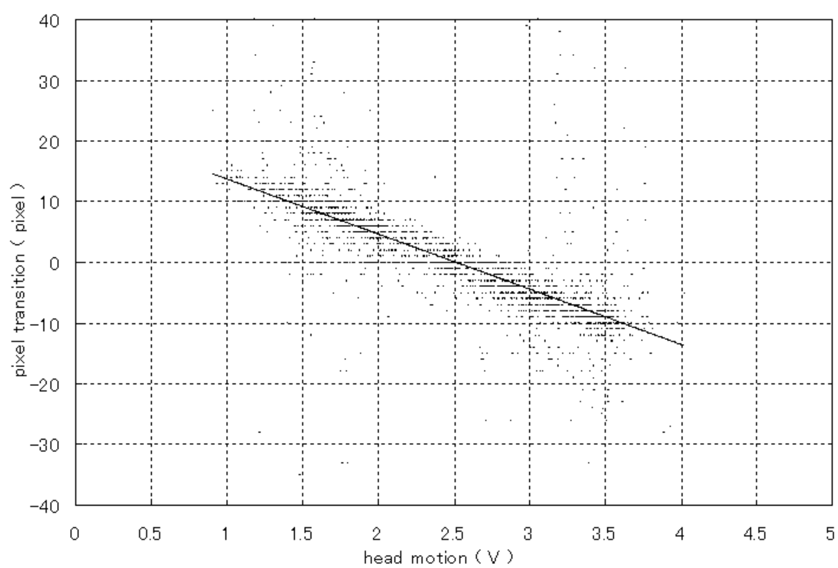


Figure 4.7 Relations between the Sensor Value and the Image Shift Value (yaw)

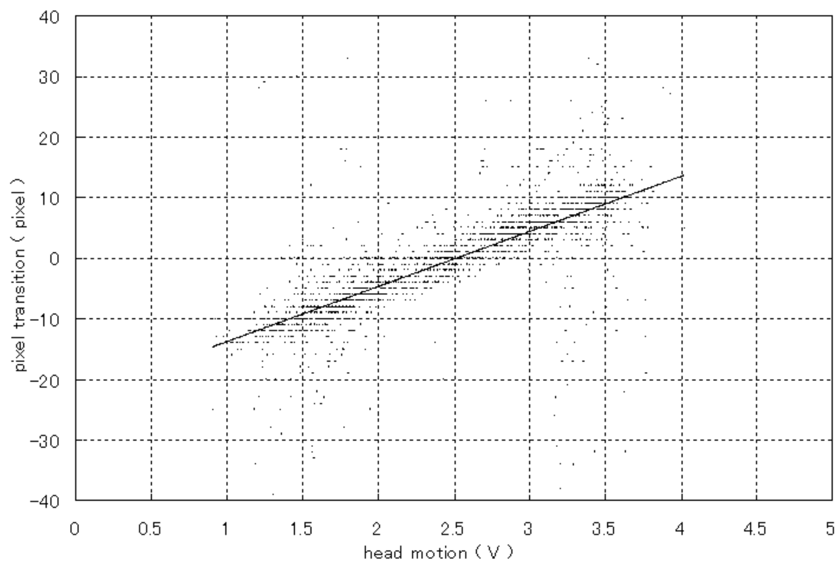


Figure 4.8 Relations between the Sensor Value and the Image Shift Value (pitch)

Figure 4.7 and Figure 4.8 show the amount relations of value transition, which occurred by head motion between the gyro sensor and an image shift. These results allow us to remove head-motion information from the video data.

- *Tracking Moving Objects*

I have employed a conventional block matching method. The matching method can detect motion areas, motion directions, and amounts of motion at the same time. This method divides an image into small blocks. Each block is an averaged and normalized RGB-value with ($I = r + g + b$, $I_r = r/I$, $I_g = g/I$, $I_b = b/I$). The method is defined as the following formulae:

$$\begin{aligned} M_{r_{i,j}}(u,v) &= \sum_{u,v}(I_r(i,j) - I_r(i+u, j+v))^2, \\ M_{g_{i,j}}(u,v) &= \sum_{u,v}(I_g(i,j) - I_g(i+u, j+v))^2, \\ M_{b_{i,j}}(u,v) &= \sum_{u,v}(I_b(i,j) - I_b(i+u, j+v))^2. \end{aligned} \tag{4.1}$$

$$M_{i,j}(u_{min}, v_{min}) = \min_{u,v} M_{r_{i,j}}(u,v) + M_{g_{i,j}}(u,v) + M_{b_{i,j}}(u,v). \tag{4.2}$$

(u_{min}, v_{min}) represents an estimated minimum motion vector value. The estimated motion block is then redifened into five simple states (up, down, left, right, and not-moving). If a motion vector is adequately small, this block is named, "not-moving."

- *Exclusion of Motion Information*

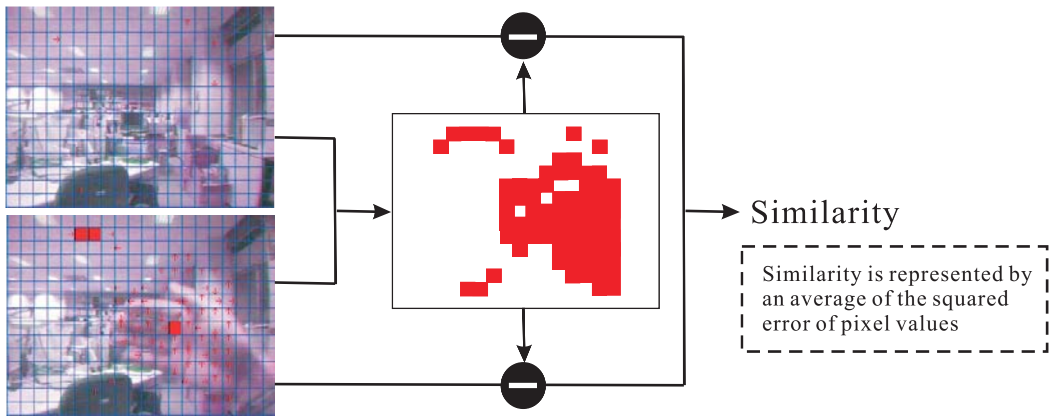


Figure 4.9. Similarity of Background as Location Information

In order to calculate the similarity of the background as location information, the process has to compare a current query image to a previous image except for motion blocks (Figure 4.9). The current query image and previous image have eigen-motion information. In order to remove mutual motion blocks in each image from target searching blocks, a motion block mask should be made. The first step of the similarity estimation compares the same address block in two images with blocks in the "not-moving" state. The second step divides a value, from which values are summed and calculated by the previous process, by the number of "not-moving" blocks.

4.2.3 Video Scene Segmentation using Motion Information

In order to segment a scene from continuously recorded video data, I employed a moving average method using two mono-axis gyro sensors. In the moving average method, continuously input gyro sensor values are added from past values in the T frames prior to the current values and the added value is divided by T. This method obtains a meta trend of captured data. The moving average method equation is as follows:

$$MA_T(t) = \frac{\sum_{i=t-T}^{t} f(i)}{T}. \tag{4.3}$$

In this study, four values are calculated by the moving average method: Two values are calculated with the yaw-axis value, and the other two values are calculated with the pitch-axis gyro value. The following three states are defined to detect scene changes:

- Stable state: Both the moving average value of the short interval (e.g. T=30) and that of the long one (e.g. T=300) are within under a certain range
- Side1 state: The moving average of the short interval is higher than that of the long one.
- Side2 state: The moving average of the long interval is higher than that of the short one.

If the difference and the amount of the value transition of gyro sensors are large, I choose a point at which to divide the scene. This step, however, does not make a new scene when a color difference between adjacent images shows under a certain threshold.

4.2.4 Real-time Video Retrieval

In this section I introduce the HySIM algorithm. The HySIM algorithm consists of two processes for high-speed video retrieval. One process includes the construction process of hybrid-space. The other process includes the process of a video scene tracking in the HySIM (Figure 4.10). In the construction of hybrid-space, the system makes a scene-segmented, video data set and constructs hybrid-space at the same time. In the video scene tracking process, the system searches for

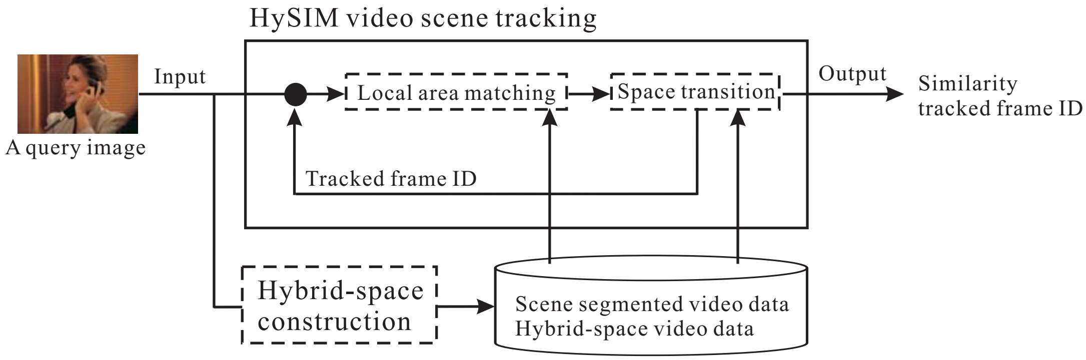


Figure 4.10. An Overview of the HySIM mechanizm

a video and then tracks a similar video using a current user's viewpoint query image from the scene-segmented video data set.

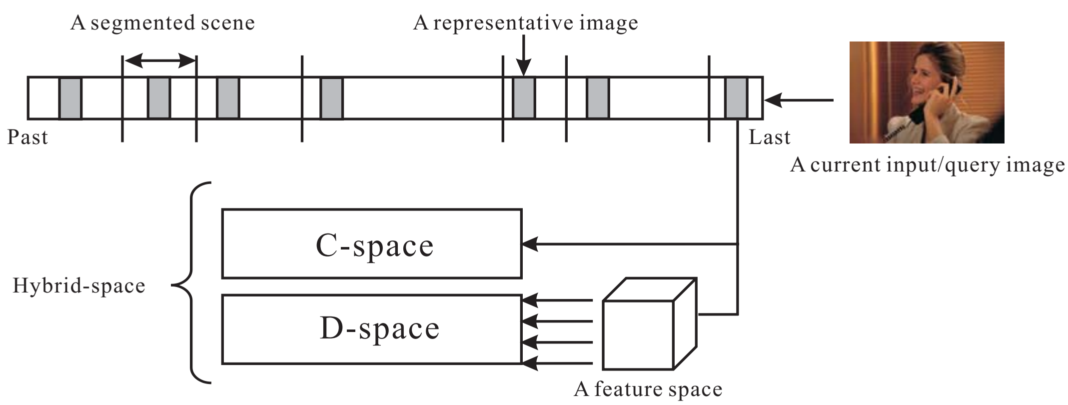


Figure 4.11. An Overview of a Hybrid-space

An overview of a hybrid-space is illustrated in Figure 4.11. The algorithm uses two spaces. These spaces consist of representative images, each of which well represent all images in a single image segment; continuously observed images are segmented so that similar images are stored in the same segmented scene. One of the spaces is a time-sequential space, which is constructed from representative images. I named this space the C-space. The other space is a feature space that is also constructed from representative images. I named this discrete space the D-space.

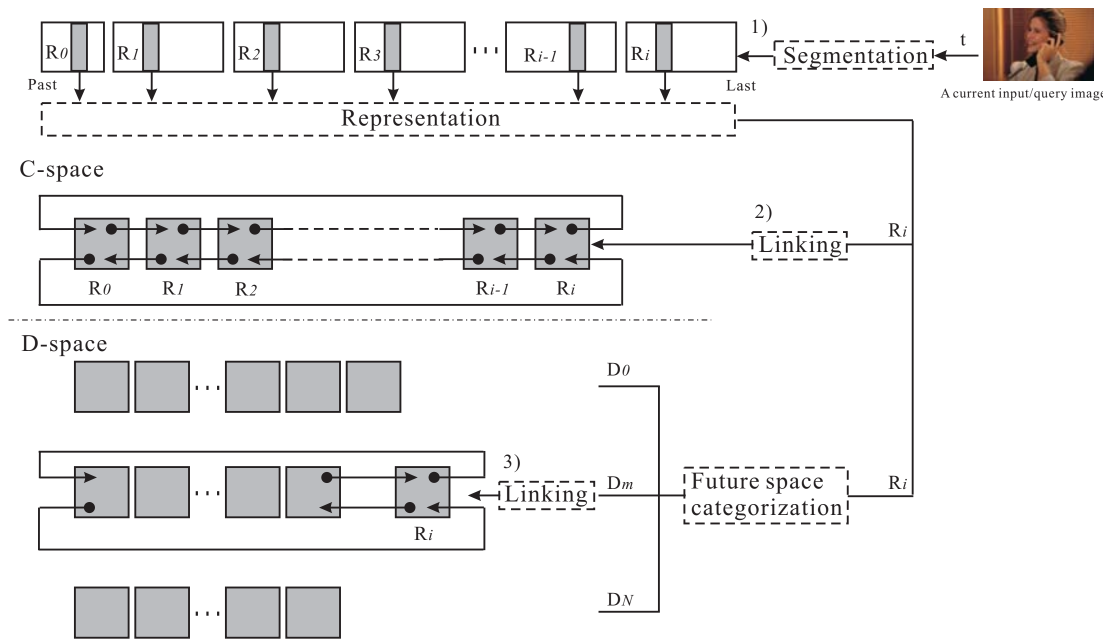


Figure 4.12. A Construction of Hybrid-space

The construction process of hybrid-space is shown in Figure 4.12: 1) Segmenting of a video scene and representation of an image from a segmented scene. 2) Linking the selected representative image to the last selected image in C-space. 3) Categorizing and linking the selected representative image to the last categorized image in the same space of D-space. In the i th segmentation and representation process, there are n numbers of a representative image candidate that include a current (t th) input/query image. A j th image in the i th segmentation process is shown $R_i(j)$. A temporal representative image R_i' is calculated by equation (4). Segmentation is performed when an error $\varepsilon_{max}(i), whichisevaluatedfromequation(5), ishigherthanathresholdTh.$

$$R_j' = \min_j \max_k(|R_i(j) - R_j(k)|), \quad (t-n \leq j,k \leq t, j \neq k), \tag{4.4}$$

$$\varepsilon_{max}(i) = \max_j(|R_i' R_i(k)|), \quad (t-n \leq k \leq t). \tag{4.5}$$

After the selection of the representative image, a hybrid-space is constructed using the representative image R_i. The C-space is reconstructed by linking representative images. R_i is linked to R_{i-1} and R_0, and R_{i-1} is unlinked from R_0. In the construction of the D-space, the down sampled Rd_i is categorized in a feature space. The representative image is then linked to the past representative image and the last image in the local space D_m of the D-space.

- *The Process of a HySIM video scene Tracking*

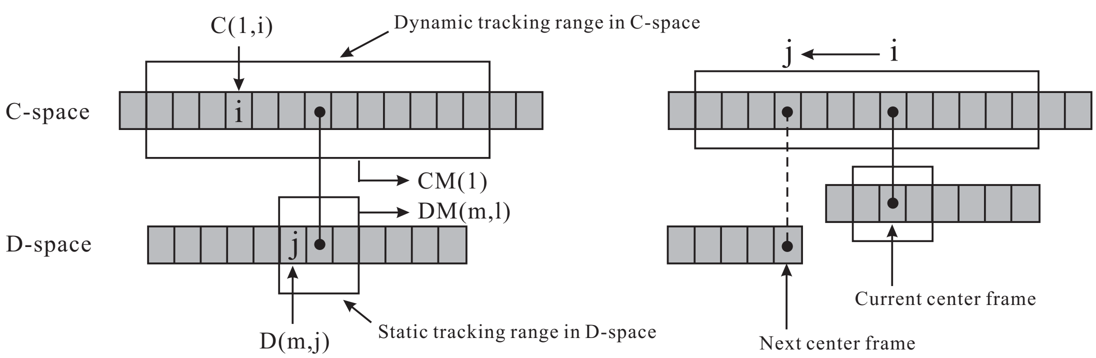


Figure 4.13. A Construction of Hybrid-space

A basic calculation method for a HySIM video scene tracking is shown in Figure 4.13. The D-space has N dimension in this thesis. A representative image is stored in both spaces. In the C-space and D-space, the tracking at a certain moment is performed within a dynamic range (CR) and a static range (DR), respectively. A value of the static tracking range in the D-space is set in advance. $D(m, j)$ represents a similarity between a representative image j and a current query image in dimension m. $D(m, j)$ is identical to $C(l, i)$. Let l denote the center frame of a tracking range in C-space and D-space. $C(l, i)$ represents a similarity between a representative image i and a current query image in the frame l. In this study, all similarities have a range from 0.0 to 1.0. In the tracking range determined by l, values of the $CM(l)$ and $DM(m, l)$ are calculated by the following equations:

$$CM(l) = \max_i (C(l,i)), \quad (l + (CR+1)/2 \leq i \leq l - (CR+1)/2), \tag{4.6}$$

$$DM(m,l) = \max_i (CM(i)). \quad (l + (CR+1)/2 \leq i \leq l - (CR+1)/2). \tag{4.7}$$

The dynamic tracking range depends on an entrrance point , similar to equation 4.8.

$$CR = \alpha \cdot A + (1-\alpha) \cdot A \cdot (1 - C(l,l))^2. \quad (\alpha, A : const) \tag{4.8}$$

Here, A represents a basic range and α represents a flexibility rate.

The next tracking step can set a new center frame by using the evaluation of the previous step and the following rule of a space transition of the tracking area.

$$l = \begin{cases} i, & \exists_i C(l,i) \equiv CM(l), l - (CR+1)/2 \leq i \leq l + (CR+1)/2, i \neq j \\ \\ j, & \exists_i CM(j) \equiv DM(m,l), l - (CR+1)/2 \leq j \leq l + (CR+1)/2 \end{cases} \tag{4.9}$$

4.2.5 Constructing a Probabilistic Histogram for Image Matching

Object recognition from color images has been widely tackled recently. In order to discriminate similar objects, a color histogram is often employed (Gevers, T. & Smeulders, W.M., 1999; Sural, S., et al., 2002), e.g, a hue color histogram. However, an image captured by a camera certainly contains additional sensor noise. Using a conventional histogram, the afore-mentioned noise reduces the object recognition rate.

In this thesis, I propose a novel probabilistic histogram, named an Integrated Probabilistic Histogram (IPH), which enables effective object recognition called . The IPH employs a probabilistic representation of color distribution in HSV color space. The IPH is made from hue data using the probabilistic behaviors of saturation and value data. The IPH is represented by the summed densities of probabilistic distributions. The IPH also has the advantage of representing a certain density if the image shows achromatic color. I conducted two experiments to evaluate the IPH; the results show its effectiveness in recognizing objects.

- *Noise Model*

The image data captured by a camera contains additional sensor noise. The noise model is shown in equation (4.10). u_{est} and σ_u are respectively the standard deviation and the estimated value.

$$\hat{u} = u_{est} \pm \sigma_u. \tag{4.10}$$

In this thesis, I define red, green, and blue values as $R, G, andB$, and also represent individual sensor noises of RGB colors, σ_R, σ_G, and σ_B in RGB space.

A standard deviation of a certain function $q(u, \ldots, w)$ is calculated by equation (4.11).

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial u}\sigma_u\right)^2 + \ldots + \left(\frac{\partial q}{\partial w}\sigma_w\right)^2}. \tag{4.11}$$

Note that $\sigma_u, \ldots, \sigma_w$ are supposed by captured images. The predicted uncertainty σ_q can be computed if $\sigma_u, \ldots, \sigma_w$ are independent, random and relatively small (Taylor, J.R., 1982). $\partial q/\partial u$ and $\partial q/\partial w$ represent the partial derivatives of q in $u, \ldots, w$.

In this thesis, I define a shape of sensor noise as the following kernel:

$$K(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-u)^2}{2\sigma^2}}. \tag{4.12}$$

In order to employ the kernel, I agreed on Gevers's kernel because photons fit well into Gaussian distribution when the average number of the counts of photons is large (Gevers, T., 2001).

- *RGB-HSV Conversion*

Values of H, S, V (and Z) in HSV color space are given by R, G, and B in RGB color space by equation (4.13) $\sim$ (4.16). In this thesis, R, G, and B are limited to $0 \leq R, G, B < \alpha$. H is also limited to $0^\circ \leq H < \beta^\circ$. The range of S and V are $0 \leq S, V < \alpha$.

Equation (4.13) gives the value of V.

$$V = max(R, G, B). \tag{4.13}$$

The minimum value of Z in R, G, and B values is given by the following equation:

$$Z = min(R, G, B). \tag{4.14}$$

Saturation value S is transformed by equation (4.15).

$$S = \alpha \left(\frac{V - Z}{V} \right). \tag{4.15}$$

H is given by equation (4.16). Suppose that $X \equiv V$ means that X ($X \in R, G, B$) corresponds with V.

$$H = \begin{cases} \beta \left(\frac{G - B}{V - Z} \right), & R \equiv V \\ \beta \left(2 + \frac{B - R}{V - Z} \right), & G \equiv V \\ \beta \left(4 + \frac{R - G}{V - Z} \right). & B \equiv V \end{cases} \tag{4.16}$$

- *Integrated Probabilistic Histogram*

The aim of the study is to be able to represent the probabilistic behavior of sensor noise on a hue color histogram. In order to realize such a representation, a calculation of standard deviation in HSV color space is required. The average and standard deviation of both V and Z, however, cannot be calculated using differentiation because of the functions of max and min.

A certain probability of $p_y(x)$ in the Gaussian distribution $y \in R, G, B$ leads to equation (4.12).

$$p_y(x) = K^y(x) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(x-u_y)^2}{2\sigma_y^2}} \tag{4.17}$$

Maximum probability $P_{X_V}(x)$ $(X \in R, G, B)$ in RGB color space is given by:

$$\begin{aligned} P_{R_V}(x) &= p_R(x)\int_{-\infty}^{x}\int_{-\infty}^{u} p_G(u)p_B(v)dvdu \\ P_{G_V}(x) &= p_G(x)\int_{-\infty}^{x}\int_{-\infty}^{u} p_B(u)p_R(v)dvdu \\ P_{B_V}(x) &= p_B(x)\int_{-\infty}^{x}\int_{-\infty}^{u} p_R(u)p_G(v)dvdu \end{aligned} \tag{4.18}$$

Equation (4.18) gives probability V in x as

$$\begin{aligned} P_V(x) &= P_{R_V}(x)(1-P_{G_V}(x))(1-P_{B_V}(x)) \\ &\quad +(1-P_{R_V}(x))P_{G_V}(x)(1-P_{B_V}(x)) \\ &\quad +(1-P_{R_V}(x))(1-P_{G_V}(x))P_{B_V}(x). \end{aligned} \tag{4.19}$$

Equation (4.19) gives an average of V as

$$u_V = \frac{\int_{-\infty}^{\infty} xP_V(x)dx}{\int_{-\infty}^{\infty} P_V(x)dx} \tag{4.20}$$

Finally, a standard deviation of V is computed by:

$$\sigma_V = \int_{-\infty}^{\infty} P_V(x)(x-u_V)^2 dx \tag{4.21}$$

The average and standard deviation of Z are also given by (4.14). Maximum probability $P_{X_Z}(x)$ $(X \in R, G, B)$ in RGB color space is given by:

$$\begin{aligned} P_{R_Z}(x) &= p_R(x)\int_{x}^{\infty}\int_{u}^{\infty} p_G(u)p_B(v)dvdu \\ P_{G_Z}(x) &= p_G(x)\int_{x}^{\infty}\int_{u}^{\infty} p_B(u)p_R(v)dvdu \\ P_{B_Z}(x) &= p_B(x)\int_{x}^{\infty}\int_{u}^{\infty} p_R(u)p_G(v)dvdu \end{aligned} \tag{4.22}$$

Equation (4.22) gives probability Z in x as

$$
\begin{aligned}
P_Z(x) &= P_{R_Z}(x)(1 - P_{G_Z}(x))(1 - P_{B_Z}(x)) \\
&\quad +(1 - P_{R_Z}(x))P_{G_Z}(x)(1 - P_{B_Z}(x)) \\
&\quad +(1 - P_{R_Z}(x))(1 - P_{G_Z}(x))P_{B_Z}(x).
\end{aligned} \tag{4.23}
$$

Equation (4.23) gives an average of Z as

$$
u_Z = \frac{\int_{-\infty}^{\infty} x P_Z(x)dx}{\int_{-\infty}^{\infty} P_Z(x)dx}. \tag{4.24}
$$

Finally, a standard deviation of Z is computed by:

$$
\sigma_Z = \int_{-\infty}^{\infty} P_Z(x)(x - u_Z)^2 dx. \tag{4.25}
$$

Substitution of (4.15) in (4.11) gives the standard deviation of S as

$$
\sigma_S = \alpha\sqrt{\frac{Z^2\sigma_V^2 + V^2\sigma_Z^2}{V^4}}. \tag{4.26}
$$

Substitution of (4.16) in (4.11) gives the standard deviation of H as

$$
\sigma_H = \begin{cases} \sigma_{RGB}(G, B), & R \equiv V \\ \sigma_{RGB}(B, R), & G \equiv V \\ \sigma_{RGB}(R, G). & B \equiv V \end{cases} \tag{4.27}
$$

where $\sigma_{RGB}(x, y)$ in (4.27) is given by:

$$
\sigma_{RGB}(x, y) = \alpha\beta\sqrt{\frac{(x - y)^2(V^2\sigma_S^2 + S^2\sigma_V^2) + S^2V^2(\sigma_x^2 + \sigma_y^2)}{V^4S^4}} \tag{4.28}
$$

Finally, the probabilistic behavior of sensor noise can be computed on a hue color histogram by the equations (4.21), (4.26), and (4.27) in HSV color space and (4.25). The probabilistic hue histogram is given by equation (4.29). The histogram also has the ability of computing achromatic color in the histogram.

$$\hat{f}_H(t) = \frac{1}{n}\sum_{i=1}^{n}\int_{h(t-1)}^{ht} K_i^H(x)dx. \quad \{t|t = 1, 2, \ldots, \frac{\beta}{h}\} \tag{4.29}$$

where n is the amount of image pixels that are selected to make the histogram, t is the parameter of bin in the histogram, and h is the width of bin.

4.3 Experiments

4.3.1 Image Matching Excluding Motion Information

- *Methods*

This experiment took place outdoors, in a hall, and in a room during the daytime. The experimental tasks were recorded four times in each place. The first image frame in the recorded video data is defined as the retrieval query. The experimental task consists of two parts. One task required the user to wave his/her hand several times. The other task required the user to turn his/her head to the left. A normal method that does not consider motion information was performed to compare with the proposed method.

- *Results*

The result is shown in Figure 4.14. In this figure, a higher similarity corresponds to a lower evaluation value. The proposed method clearly shows a higher similarity than the normal method in the hand waving section. In comparison, eliminating noises caused by turning head was less effective than that by moving a hand in the method.

Table 4.1 illustrates the precision and recall rates. The precision rate represents the rate of retrieved data for all correct data. The recall rate represents the rate of correct data for all retrieval data. Therefore, the proposed method is well suited to retrieving similar location-based images because the relevance rate of the proposed method performed 1.3 times better than the normal method.

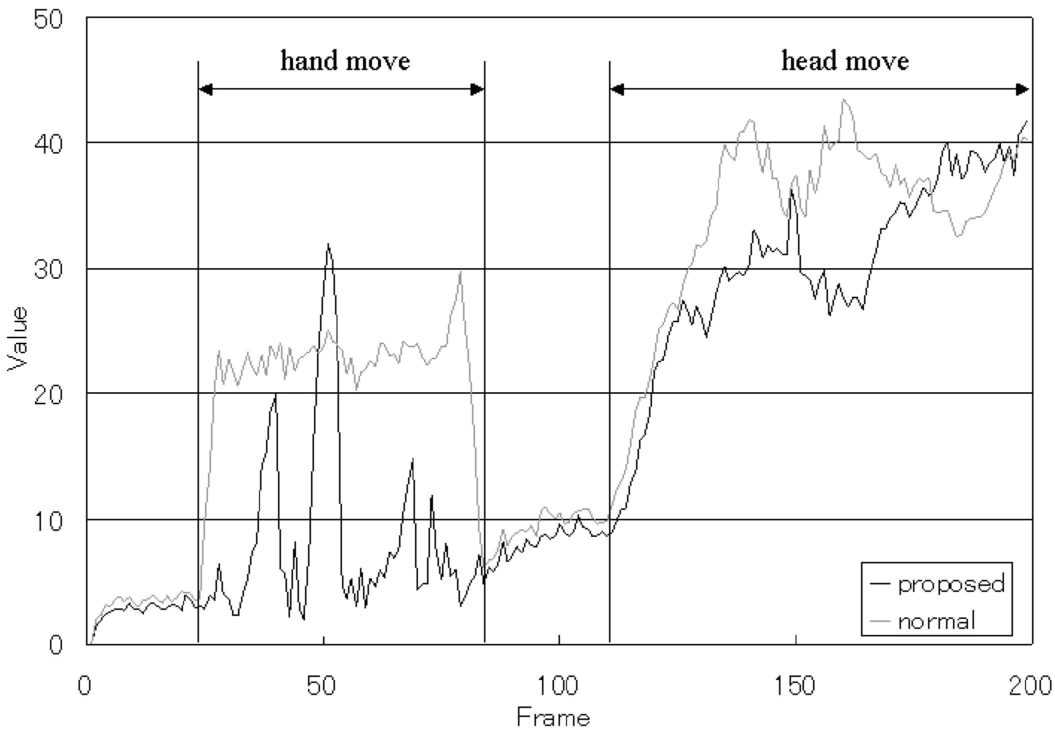


Figure 4.14. Evaluation and Targeting of Video Similarity between Base Image

Table 4.1. Precision and Recall Rate

	Precision		Recall	
	Proposed	Normal	Proposed	Normal
Outdoor	90%	54%	98%	96%
Hall	97%	56%	92%	90%
Room	88%	57%	97%	96%
Total	92%	56%	96%	94%

4.3.2 Video Scene Segmentation using Motion Information

- *Methods*

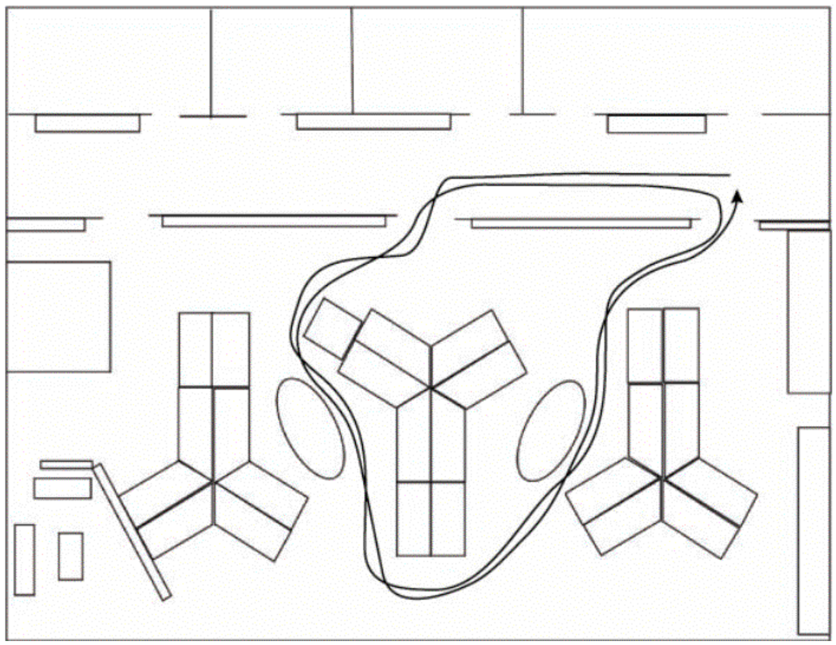


Figure 4.15. The Walking Route of the Test Subject

Both image and gyro data, which consist of 5584 frames (191.36 seconds, 29.19 frame/second), were recorded for evaluation. A subject walked around a certain route two times in the laboratory I belong to (Figure 4.15). I set intervals of the moving average as $T = 30$ and $T = 300$. This process limits the minimum scene length to 30 frames.

- *Results*

This remarkable result of the experiment is shown in Figure 4.16. The upper lines in the figure are the moving average values. The lower line shows the detected scene changes. These scene changes took place 9 times in the figure. The following results (5584 frames) were obtained: 41 total scenes were constructed

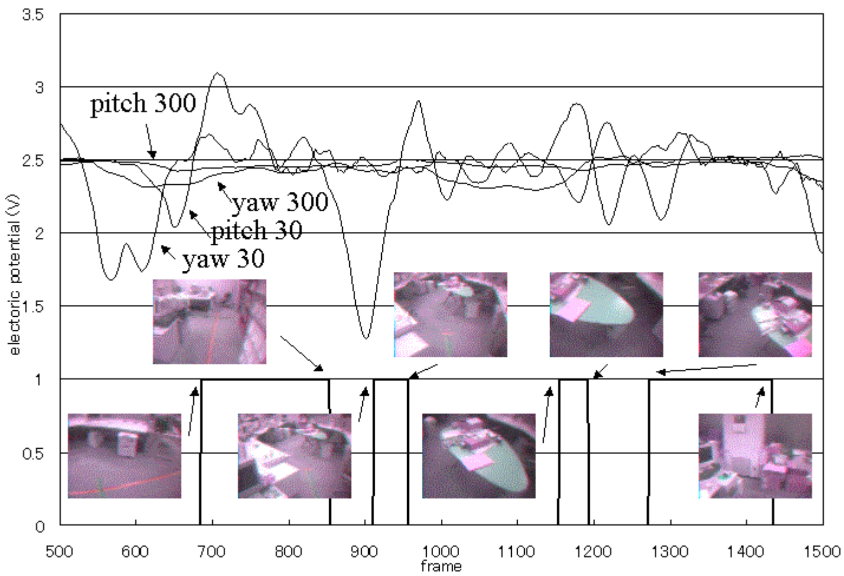


Figure 4.16. Making Scene Changes using Two Mono-axis Gyro Sensors

from a video data set in the experiment. The average length of the constructed scenes was 119.51frames (4.09 seconds). The minimum and the maximum length of these scene were 31 frames (1.06 seconds) and 756 frames (25.90 seconds), respectively. The minimum scene was made when the subject walked quckly through a narrow path. The maximum scene change was constructed on a long, straight, and undiversified hallway. From this experimental result, I conclude that the user's motion information can be used to construct scene changes on a wearable computer environment.

4.3.3 Real-time Video Retrieval

- *Methods*

The following experiments were conducted to evaluate the video retrieval performance of the HySIM algorithm. In these experiments, the HySIM algorithm is compared with a full-search algorithm, which uses only representative images. I have evaluated retrieval speed and retrieval accuracy.

In these experiments, one test subject wore a wearable computer system, and continuously walked around for approximately an hour in a building with three halls and two laboratory rooms. Input query images were captured frame by frame. I used the same images for the query data set as for the recorded data set. 100,000 images were captured for approximately one hour. The parameter α is set at 50 in this experiment. The parameter DR value is equal to 15 per trial.

\- *Results*

Table 4.2. A comparison between the full-search and the HySIM search

Data (frame)	R-frame (frame)	Full (sec)	HySIM (sec)	Accuracy rate (%)
10,000	1257	.0106	.0021	69.48
20,000	2254	.0189	.0022	69.90
30,000	6235	.0515	.0031	34.07
40,000	10138	.0835	.0035	35.73
50,000	11142	.0920	.0033	34.70
60,000	12568	.1037	.0033	31.93
70,000	13724	.1138	.0031	34.69
80,000	15313	.1268	.0031	41.92
90,000	16520	.1374	.0030	37.61
100,000	21039	.1744	.0033	30.33

Table 2 shows the performance result of these experiments, in which the data column shows the number of images. The R-frame illustrates the amount of representative images in a trial. Both full and HySIM methods show the video data processing times required to retrieve a similar image with an input query image, respectively. The accuracy rate is calculated by the number of frames, which have a similarity with over the threshold Th (set at 0.95 in the experiment) in the scene segmentation, or as the best similarity in all recorded images.

The processing time of the full algorithm shows a linear increase. The processing time of HySIM increases approximately 1.57 times when the data of 10,000 frames increases to 100,000 frames. The processing time of HySIM is 52.20 times as fast as that of the full algorithm in the data of 100,000 frames.

4.3.4 Constructing a Probabilistic Histogram for Image Matching

\- *Methods*

I conducted two experiments to evaluate the quality of the IPH. In the experiment, the IPH was compared with a conventional/simple hue histogram. Note that parameters α and β were set as $\alpha = 255$, and $\beta = 359$. R, G, B, H, S, V, and Z were over 1 interval. The experiments were conducted on a cylindrical target having a gradual circumferential pattern change of saturation and value. The difference between the source histogram and a target histogram was computed using the sum of absolute differences (SAD) in each histogram. The conditions of the experiments are shown as follows:

C1: A comparison between the same posture images

C2: A comparison between different posture images

Figure 4.17 depicts 12 postures of the cylindrical target. 10 images were continuously captured in each posture over 30° intervals. The evaluated cameras all got 120 images in the same environment condition.

I employed three types of color cameras for the experiments. Former two cameras are developed for recognizing objects. One is a pre-production sample camera, which is named "*EVIS*," by SONY (Sugiyama, T, et al., 2002). The device alternately captures color image and IR image using only one photo acceptance unit. Another camera termed "*ObjectCam*" consists of a color camera and an infrared camera clamped at the same posture across a beam splitter (Ueoka, T., et al., 2003). A frame of the image captured by the *ObjectCam* consists of both a color field and an infra-red (IR) field image. Last one is a USB and CMOS-CCD

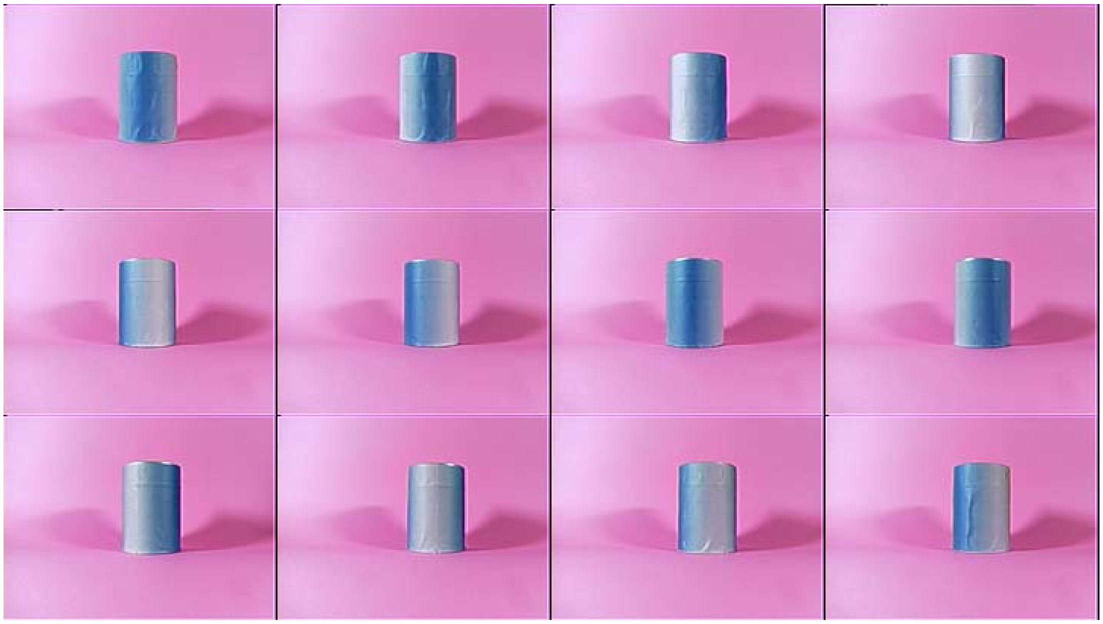
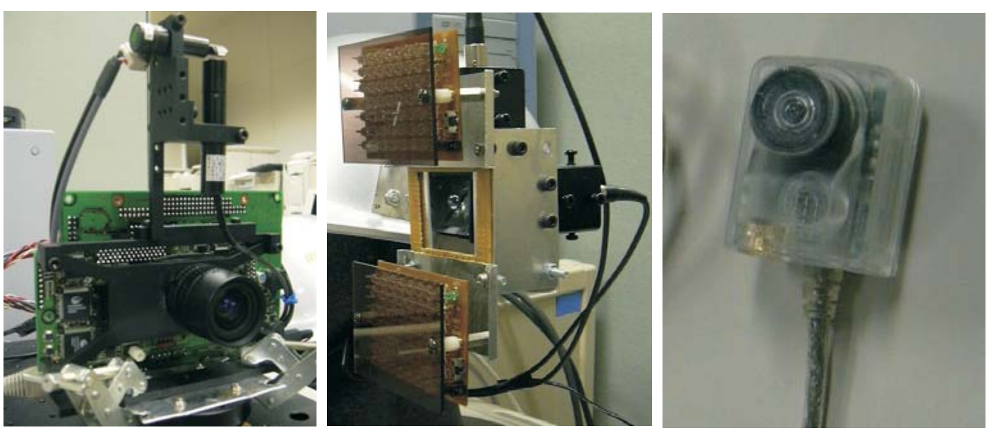


Figure 4.17 A cylinder target changing satulation and value along with the circumferential (Top：0°, 30°, 60°, 90°. Middle：120°, 150°, 180°, 210°, Bottom：240°, 270°, 300°, 330°.)



(a) EVIS (b) ObjectCam (c) USB CCD

Figure 4.18. Camera Devices Used for the Experiments

camera produced by I-O DATA. Figure 4.18 represents camera devices used for the experiments. Measured sensor conditions are shown as follows:

Cam1

TYPE : SONY Entertainment Vision Sensor (EVIS)

SD : $\sigma_R = 1.01,\ \ \sigma_G = 0.86,\ \ \sigma_B = 1.04$

Cam2

TYPE : ELMO UN43H (ObjectCam)

SD : $\sigma_R = 3.52,\ \ \sigma_G = 2.23,\ \ \sigma_B = 3.20$

Cam3

TYPE : I-O DATA USB-CCD

SD : $\sigma_R = 4.19,\ \ \sigma_G = 2.35,\ \ \sigma_B = 3.41$

- *Results*

Figure 4.19 shows the average SAD ratio of the IPH on the conventional hue histogram in condition C1. The vertical axis represents an evaluated value where the average error of SAD on the IPH was divided by the average error of SAD on the conventional histogram. In the *Cam1* and *Cam2* condition, each total average error showed less than 0.4 points. In all camera conditions, the total summed average errors were 0.30 in *Cam1*, 0.32 in *Cam2*, and 0.44 in *Cam3*.

Figure 4.20 illustrates the standard deviation SAD ratio on the IPH to the conventional hue histogram in condition C1. The vertical axis represents an evaluated value as the standard deviation error of SAD on the IPH, which was divided by the standard deviation error of SAD on the conventional histogram. In all camera conditions, the total summed standard deviation errors were 0.57 in *Cam1*, 0.81 in *Cam2*, and 1.18 in *Cam3*.

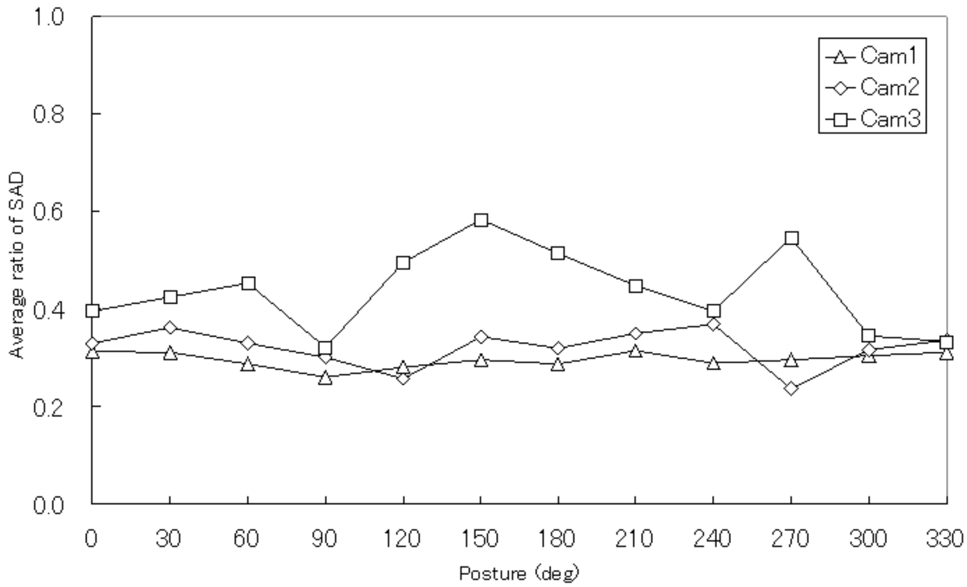


Figure 4.19. Average ratio of SAD between the same postures

Figure 4.21 describes the average SAD ratio on the IPH to the conventional hue histogram in condition C2. The source posture was set at 0° posture. The vertical axis represents an evaluated value where the average error of SAD on the IPH was divided by the average error of SAD on the conventional histogram. In all camera conditions, the total summed average errors were 1.90 in *Cam1*, 1.13 in *Cam2*, and 1.29 in *Cam3*.

Figure 4.22 depicts the standard deviation SAD ratio on the IPH to the conventional hue histogram in condition C1. The source posture was set at 0° posture. The vertical axis represents an evaluated value where the standard deviation error of SAD on the IPH was divided by a standard deviation error of SAD of the conventional histogram. In all camera conditions, the total summed standard deviation errors were 0.86 in *Cam1*, 0.98 in *Cam2*, and 0.84 in *Cam3*.

In condition C1, the total summed average errors were 0.30 in *Cam1*, 0.32 in *Cam2*, and 0.44 in *Cam3*. The results show that the IPH depresses the average error of SAD among the same posture images less than in the conventional histogram. In the C2 condition of the experiments, the total summed average errors were 1.90 in *Cam1*, 1.13 in *Cam2*, and 1.29 in *Cam3*. The two results show the effective suppression of the average error of SAD among the different posture

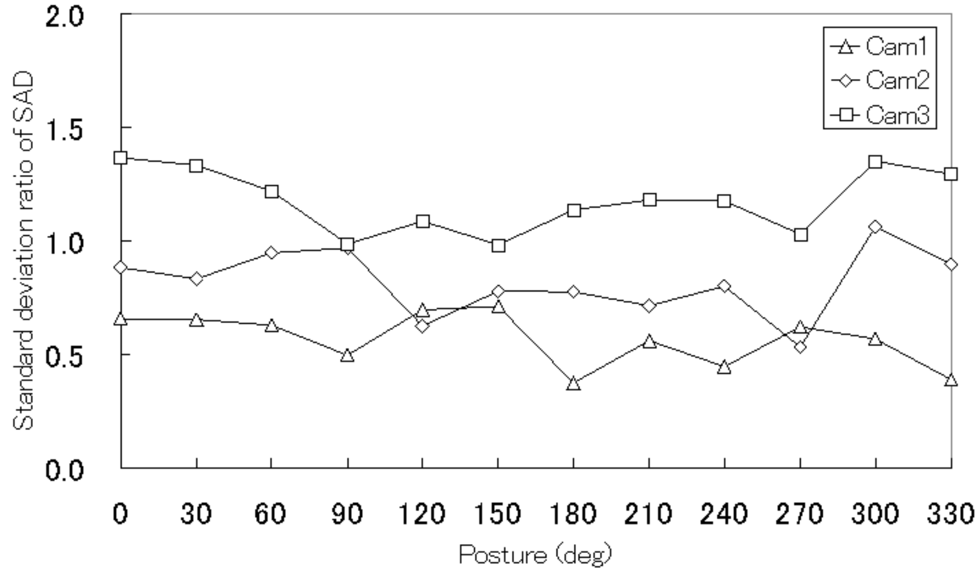


Figure 4.20. Standard deviation ratio of SAD between the same postures

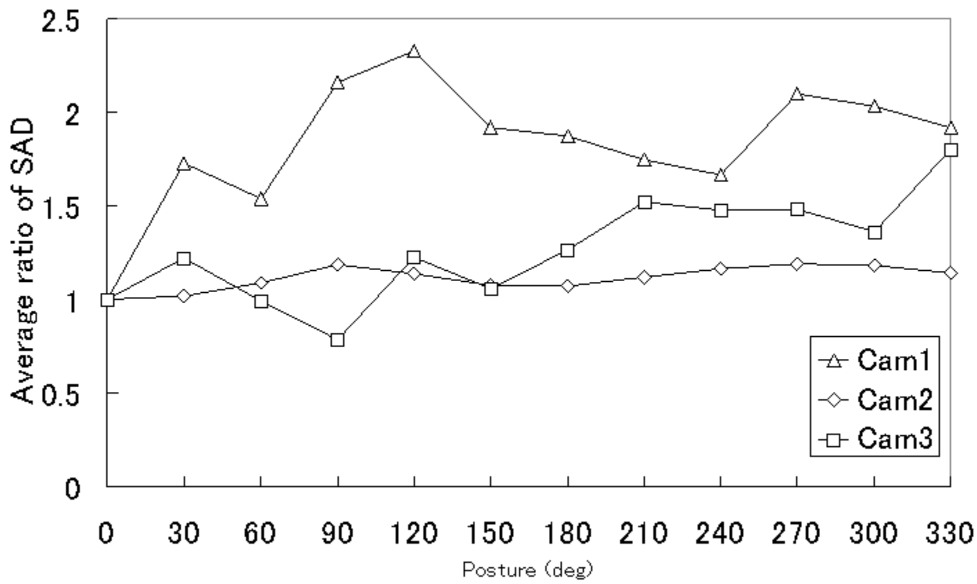


Figure 4.21. Average ratio of SAD between different postures

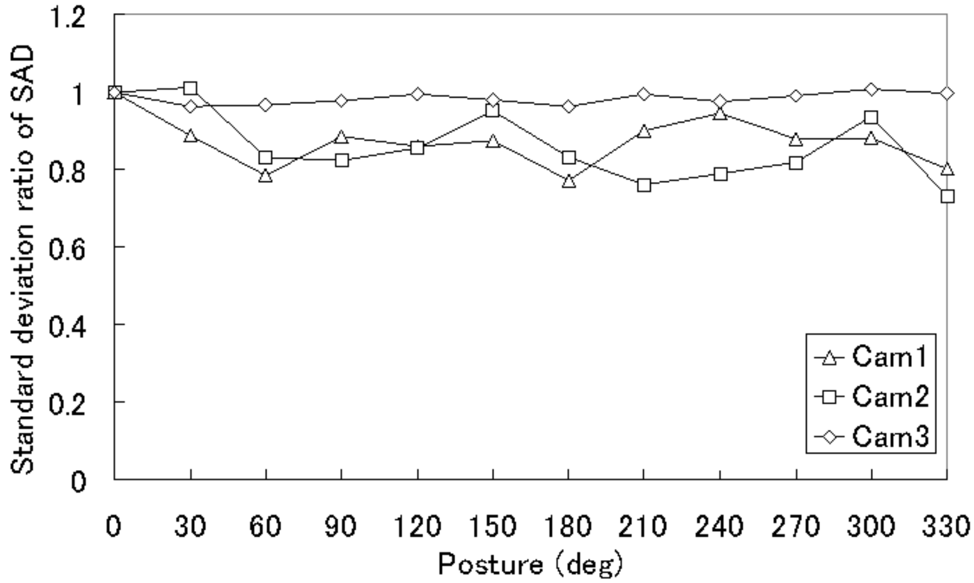


Figure 4.22. Standard deviation ratio of SAD between different postures

images. These contrasting results between C1 and C2 give us considerable benefit in discriminating objects on a vision-based application.

The results of the C1 condition experiment are 0.57 in *Cam1*, 0.81 in *Cam2*, and 1.18 in *Cam3* in the standard deviation SAD ratio of the IPH to the conventional histogram. Also, the results of the C2 are 0.86 in *Cam1*, 0.98 in *Cam2*, and 0.84 in *Cam3*. In those results, depressions of standard deviation are shown, except for *Cam3*. The results show that the IPH provides the system with stability for object recognition.

4.4 Related Works

The following studies use a video and a stand-alone type wearable system. Clarkson et al. proposed a method for segmenting sequential audio and video data (Clarkson, B. Pentland, A., 1999; Clarkson, B. Pentland, A., 2000). This method can recognize situations of abstract location. However, this method cannot directly retrieve previous associable video data for a user who wants to know detailed location information. Aoki, H., et al. (1999) proposed a real-time personal positioning method with video data. This system is suitable for navigating a user to an unfamiliar location. However, selecting similar video images quickly with a continuously recorded video is difficult because an offline training process is used. Hoisko, J. (2000) proposed a locaion-based memory triggering system using IR tags attached to the room. However, the system is difficult to choose retrieved videos for keeping user's workload low when the camera detect many IR tags at the same time.

4.5 Summary

I introduced three methods for supporting a location-triggered user's recollection in the *Residual Memory* module. First, I proposed the stable image matching method concerning head movement and moving objects in a scene. I employed two mono-axis gyro sensors to track a two-axis user's head movement. Also, I proposed the video scene segmentation method to create comprehensible video scenes for the user. I introduced the moving average method by using two mono-axis gyro sensors. In addition, I proposed a real-time video retrieval method named HySIM. I introduced two different types of feature spaces: a time-sequential space, and an image-feature space with a similar color feature value. Last, I proposed a probabilistic HSV color histogram called IPH to reduce sensor noize for an image matching.

A further direction of this study will be to develop a faster, more stabilized, and more efficient video retrieval method. Additional types and numbers of perceptual sensors should be attached to achieve this goal. Shortcoming of this study, however, have remained in each implementation techniques.

In the location information extracting method using two mono-axis gyro sensors, the following two issues are left. First, the system should continusously track moving objects that are in captured images when these objects become a state of rest in the scene. Furthermore, the proposed method cannot recognize any "moving" objects that are temporarily standing at the same place. Next, the system should track user's head movement in higher degrees of freedom. The proposed method employed two mono-axis gyro sensors for detecting motions in yow and pitch axis. This method should have roll axis and all axis of linear motion for perfectly extracting location information.

A shortcoming of the HySIM method is shown in low precision rate of video retrieval although the method experientially gives enough performance to retrieve and display videos under the condition of 100,000 frames of recorded images in the precision rate of the experiment. One of reasons might be got by a serial tracking process itself because there are always some candidate video segments for a retrieval in a tracking process. Hence, the tracking method should track some candidates at the same step. And then the method compares with these candidates for picking up one candidate as a starting point of the next step.

The IPH method has had two troubles. One is in a parameter setting of a standard deviation of RGB color. In this study, the standard deviations of RGB was measured in an image totally. However, photon noise in each photo acceptance unit show respectively different characteristics. I guess that the total parameter setting prevents the comparison precision of the IPH from increasing.

Chapter 5

Nice2CU - Physical Human -

Main issues of this chapter are to facilitate a real world-oriented management method. In the method, it is important for a user to easily memorize and recall the information related to a person when the user meets the person. The following technical topics are important to solve how a human-triggered augmented memory should be managed in the real world.

- modeling information related to human
- designing module operations and implementing real world-oriented interfaces for reducing a management cost
- distinguishing friends a user knows and modeling dynamic human relations to recommend augmented memories for a rich conversation

5.1 Design

This chapter represents an example of the operational category of "Memory Retrieval"and "Memory Editing." in Chapter 3. Also, this chapter shows an example of the associable category of the "Physical Human" element mentioned in Chapter 3. Figure 5.1 depicts a model of a human-triggered memory retrieval. A user pays an attention to a certain person. The user tends to recall the person's

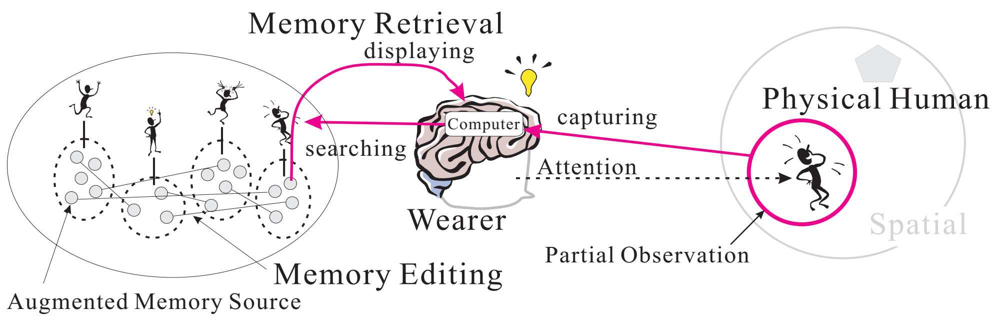


Figure 5.1. Human-triggered Augmented Memory Retrieval

information, e.g., name, affiliation, messages, experiences. A human-triggered augmented memory module captures a users viewpoint image, searches an augmented memory source associated with the target person, and then the module shows the retrieved augmented memory source to the user. The user augments his/her memory by recalling the target's information he/she wants to recall or he/she forgets.

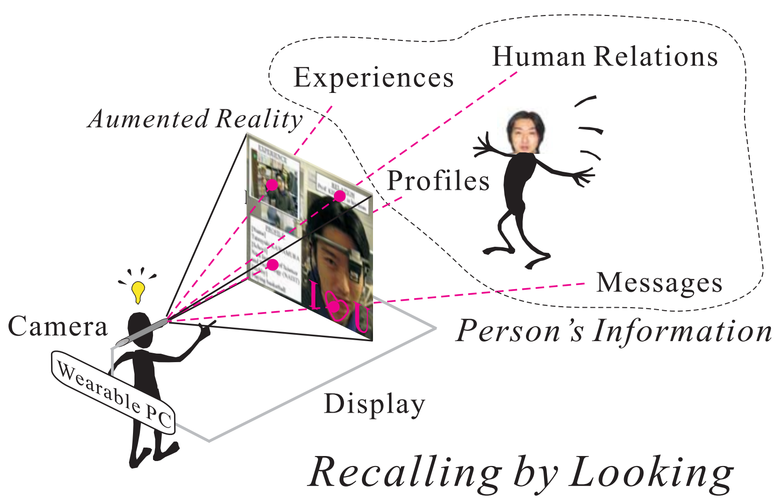


Figure 5.2. Concept of Nice2CU

I have designed a prototype system named Nice2CU (Nice to Communicate among Users). A main issue of this study is how to manage person's information in the real world. Figure 5.2 shows a concept of the *Nice2CU* that enables a wearer to recall the person's information by looking at the target person. I have proposed an easy registration method and an automatic update method. The aim of these methods is to enable the system to give a wearer the latest information of a target person who stands in front of the wearer. Realizing these methods, I also have defined a data set named a Personal Profile Set (PPS) composed of a certification data and a profile data. The easy registration method allows the wearer to identify the target person. The update method is an automatic update that can send the latest PPS to an entire group of wearers registering the target person.

In this chapter, I also propose a rating method for clarifying human relations to recommend an augmented video memory, which will enable a user to recall information he/she wants to know, using human profile data and meeting logs. I termed this set of data and logs is called a "Dynamic Profile Network." A wearer views a video of his/her interest while talking with a target person who is standing in front of him/her. Viewing this video is helpful to the person when he/she cannot recall what topic is good for the context fo a conversation. Also, selecting an appropriate video from a huge video data set recorded previously is difficult for the wearer. I categorize the wearer's friends into three types of friends. Additionally, I propose rating rules for evaluating human relations, and divide the rating rules into social rules and personal rules.

An implementation of the above methods can be achieved by a "Card (for registering) and Mirror (for updating)" interface. The "Card and Mirror" system is one of wearable systems for augment-able reality (Rkimoto, J., et al., 1998). The system employs an RFID business card and a magic mirror interface. The system provides the wearer with a direct referring operation and an indirect referring operation to refer to the target person's information. In the direct referring operation, the system employs a face recognition technique similar to Farringdon, J. Oni, Y., (2000) and Kato, T., et al. (2002a; 2002b) while the wearer directly meets the target person. In the indirect referring operation, the wearer can indirectly refer to the target person's information via the card.

Figure 5.3 shows an actual case of saying "Nice to meet you" and exchanging a business card. Such case is normally seen in business situation in Japan, and

Figure 5.3. A Case of “Nice to meet you”

represents a simple reciprocal self-disclosure. This case is realized by using a Card interface for easy registration. A business card is attached/implanted with an RFID tag. The RFID tag records the URL of the target person. By using the card, the recipient can access the PPS of the target person.

5.1.1 Conceptual Design

The aim of Nice2CU is to annotate a wearer's view with information about a person who stands in front of the wearer. The Nice2CU enables the wearer to manage and refer to the information of the target person in the real world. Information about a person for augmenting wearer's memory is defined as the following four types (Figure 5.4):

PROFILES includes a current information set about the person. (e.g. person's face, name, sex, birthday, blood type, birthplace, current address, and affiliation.)

EXPERIENCES contains an event memorized by a certain person, and includes the context of what happens in the event that was experienced together by a recipient and a target person.

MESSAGES includes temporary memorandum like a “post-it” for delivering the message to the target person. For instance, the recipient suddenly

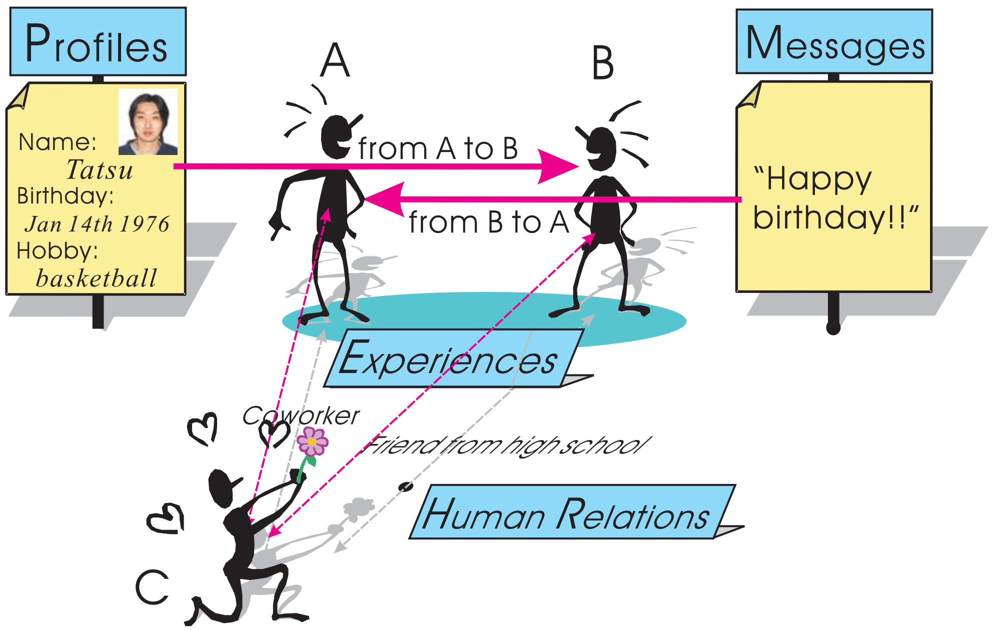


Figure 5.4. Four Types of Elements about Person's Information

remember that he/she wants to repay money to the target person when the recipient accidentally runs into this target person.

HUMAN RELATIONS consists of the above elements: PROFILES, EXPERIENCES, and MESSAGE. For example, the target person was the recipient former boss when you were in college.

In addition to the Card and Mirror interface, the Nice2CU system employs the after-mentioned *Ubiquitous Memories* that associates video data with a real world object. Figure 5.5 depicts how the wearer can use the experiences of a certain person. The wearer can associate video data recorded as episodic augmented memory with a certain person by an indirect referring operation, touching the Card.

Figure 5.5. Associating Experiences with a Person

5.2 Implementation

5.2.1 System

- *Wearable Equipment for the Nice2CU system*

Figure 5.6. The Nice2CU Equipment

Figure 5.6 shows the equipment worm with Nice2CU. The user wears the SHIMADZU HMD, which is named DataGlass2, remodeled by the Hitachi Zosen

Table 5.1. An RFID Tag Information

Purpose	Object Identification Number	Data
Address for strage	SRN	URL

(left). The HMD has a USB camera attached to the center of display. The user also wears an RFID tag reader/writer produced by OMRON to his/her wrist. In order to control the system, the wearer attaches RFID operation tags to the opposite side of wrist from the RFID tag reader/writer. The wearer carries a wearable computer on his/her hip. The RFID device can immediately read an RFID tag data when the device come close to the tag. The entire system connected to the World Wide Web via a wireless LAN.

- *A Businesss Card Attached to an RFID tag*

I assume that an RFID tag is attached to/implanted in each business card. I have employed a short-range type RFID system for 1) identifying each real world object, and for 2) controlling the states of the system. The readable range of the RFID strongly depends on the RFID tag size. A small size tag of less than 1cm allows the reader to read the tag. Data can also be retrieved from a large tag of about 3cm. Figure 5.7 depicts a sample image of an RFID tag attached to a business card. The Nice2CU system gets information from an IC of the RFID tag. Table 5.1 describes a tag data protocol to maange augmented memory. The data concerns identifying a certain business card and the card's owner. I have employed a Serial Number (SRN), which is unique to each RFID tag before shipping, as the card identification number.

- *The Mirror Interface*

Figure 5.8 shows the mirror equipment for the "Card and Mirror" intarface. The mirror has a IEEE1394 camera set at the back of a magic mirror (right). The mirror also has an RFID tag reader/writer produced by OMRON. In order to control the system, the wearer attaches RFID operation tags to the wrist. The

Figure 5.7. An RFID Tag Attached to a Business Card

RFID device can immediately read an RFID tag data when the device come close to the tag. The entire system connected to the World Wide Web via a wireless LAN.

To implement an automatic update, the system employs the mirror equipment. The target updates his/her own PPS everytime he/she stands in front of the mirror. The mirrir sends to the target's *Nice2CU*, and then the target's sytem sends his/her latest PPS to the entire group of registered recipients. The recipient's system can recognize the face of the target person using his/her latest information of the certification data.

- *A System Configuration Diagram*

Figure 5.9 shows the system configuration of *Nice2CU*. This system is composed of a wearable system and a mirror system for a wearer. The wearable computer plays the role of a client system. The core process is Nice2CU Control on the client system. The Video Buffer Control temporally manages a video memory before it is stored in a database. The Ring Buffer enables the wearer to record a video when he/she meets with other target person. The system can

Figure 5.8. The Mirror Equipment

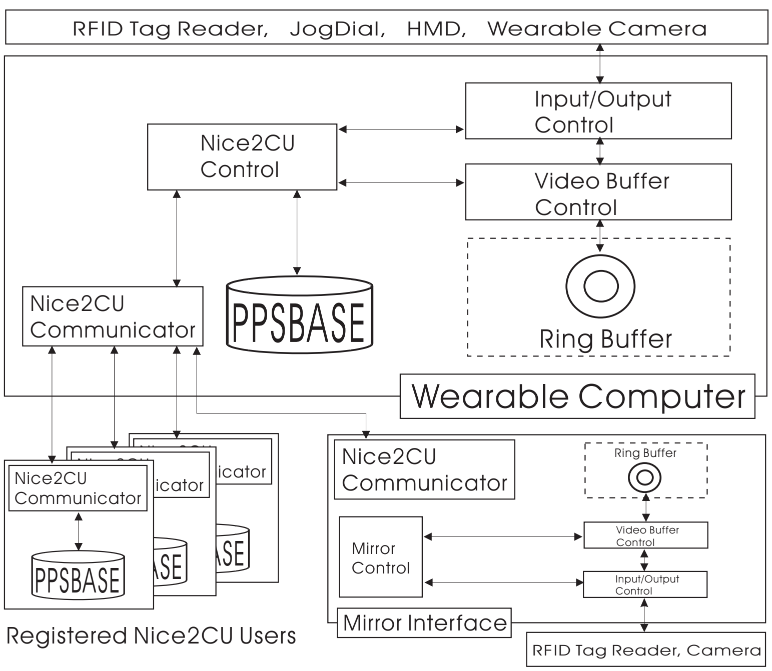


Figure 5.9. Nice2CU System Configuration

record video while the target stands in front of the wearer. This system has a Personal Profile Set dababase (PPSBASE). The mirror interface also shows similar configuration to the wearable system. The mirror interface send a new PPS to the wearer via WWW.

- *Operation Examples of the Nice2CU*

Figure 5.10 represents an example of a transition from “Nice to meet you” to “Nice to see you.” on the *Nice2CU*. (i) A user read an article in a room. (ii) The user once met with a target person. (ii) The target gave the user his/her own business card. (iv) When the user touched the card, the *Nice2CU* read information of the RFID tag attached to the card. (v) The module annotated the target’s profile, and continuously recorded video while the module identified the target. (vi) The module automatically replays the target’s profile and the previous recorded video when the user meets with the target again.

Figure 5.11 represents an example of an operation for associating a video message with a card. (i) The user read a certain comic. He/she was interested in the comic, and remembered a certain friend who liked reading comics. (ii) The user picked up a card given by the target friend. (iii) The user touched the card, (iv) and then the *Nice2CU* started recording a video to leave a message. (v) The module automatically displays the target’s profile and replays the left message for the target when the user meets with the target. (iv) Finally, the user can recommend the comic to the target.

5.2.2 Personal Profile Set

A Personal Profile Set (PPS) is composed of a certification data and a profile data. The certification data is used for identifying the target who stands in front of the recipient. The profile data is a list of the target’s profiles.

- *Certification Data*

The certification data is used to identify a person who meets the wearer. The *Nice2CU* could employ face image, a feature from the image, and signal pattern

(i) (ii) (iii) (iv) (v) (vi)

Figure 5.10. A Easy Registration through Video Replay Procedure

Figure 5.11. A Message Linking through Video Replay Procedure

to identify the person. Conventional study of the identification the person who is standing in front of a camera suppose that he/she does not change his/her own hair style. His/her costume can be employed to identify him/her during a certain day, so that dairy updating of his/her own certification data is important for a more robust certification in the real world.

- *Profile Data*

Profile data is composed of personal profile data and meeting logs. The profile data has a merit for computing human relations. Human relations can exist with a fact of "meeting others." In order to analyze more detailed aspect of the human relations, the module must have his/her background and historical data. Here, there are fundamental elements are needed to define the personal profile data and meeting logs as follows:

Date: year, month, day, hour, minute, and second

Address: town, city, prefecture/state, country, ZIP code

Then, each information is constructed as follows:

- **Personal Profile Data** means additional identical information of a person, e.g., name, birthday, sex, current address, affiliation, etc. These information are more exactly defined as the following elements:

  Inborn Facts:

  - **Sex:** male/female
  - **Birthday (Date):** year, month, day, hour, minute, second
  - **Birthplace (Address):** town, city, prefecture/state, country
  - **Bloodtype:** A/B/O/AB, (RH+/RH-)

  Name:

  - **From (Date):** year, month, day ,hour, minute, second

- **To (Date):** year, month, day ,hour, minute, second
- **Name:** first, middle, last

Address:

- **From (Date):** year, month, day ,hour, minute, second
- **To (Date):** year, month, day, hour, minute, second
- **Address (Address):** town, city, prefecture/state, country

Affiliation:

- **From (Date):** year, month, day ,hour, minute, second
- **To (Date):** year, month, day ,hour, minute, second
- **Affiliation Name:**
- **Affiliation Address:** town, city, prefecture/state, country

Hobby:

- **From (Date):** year, month, day ,hour, minute, second
- **To (Date):** year, month, day ,hour, minute, second
- **Category:**
- **Content:**

- **Meeting Logs** represents dynamic information of human relation between a recipient and a target persons. In terms of the *Nice2CU*, the module captures a conversation scene video in a meeting. The data of "Who does the recipient meet with" is also important things to compute not only human relation between the recipient and the target and also human relations among the target and other friends of the recipient. The following definitions show the meeting logs in this work:

  Record Log:

- **From (Date):** year, month, day ,hour, minute, second
- **To (Date):** year, month, day ,hour, minute, second
- **Owner:** user ID
- **Data Type:** Experience/Message
- **Filename:** filename

Replay Log:

- **Last Replay (Date):** year, month, day ,hour, minute, second
- **Replay Times:** amount of replay times
- **Filename:** filename

Friends Log:

- **User ID:** friend's user ID
- **Friend Type:** type

5.2.3 Easy Registration and Automatic Update

- *Easy Registration*

Figure 5.12 illustrates a registration procedure of the target person. In the Nice2CU system, there are three steps for the registration procedure of the PPS: 1) The target gives the recipient a target's own URL(target), which shows the address of the target's system. The recipient then registers the URL(target). 2) In order to request the target's PPS, the recipient's system sends the target's system its own URL(recipient). The target's system then registers the URL(recipient) in a registered list to update a registering recipient's PPS base. 3) The target's system uploads the PPS to the recipient's system.

- *Automatic Update*

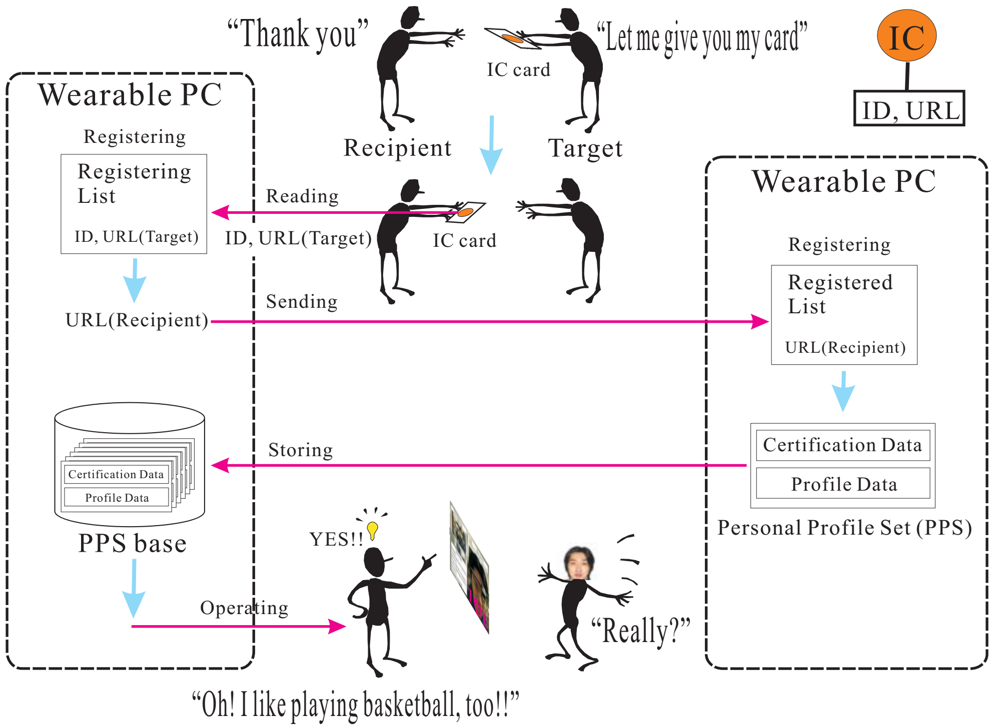


Figure 5.12. Registration of a PPS

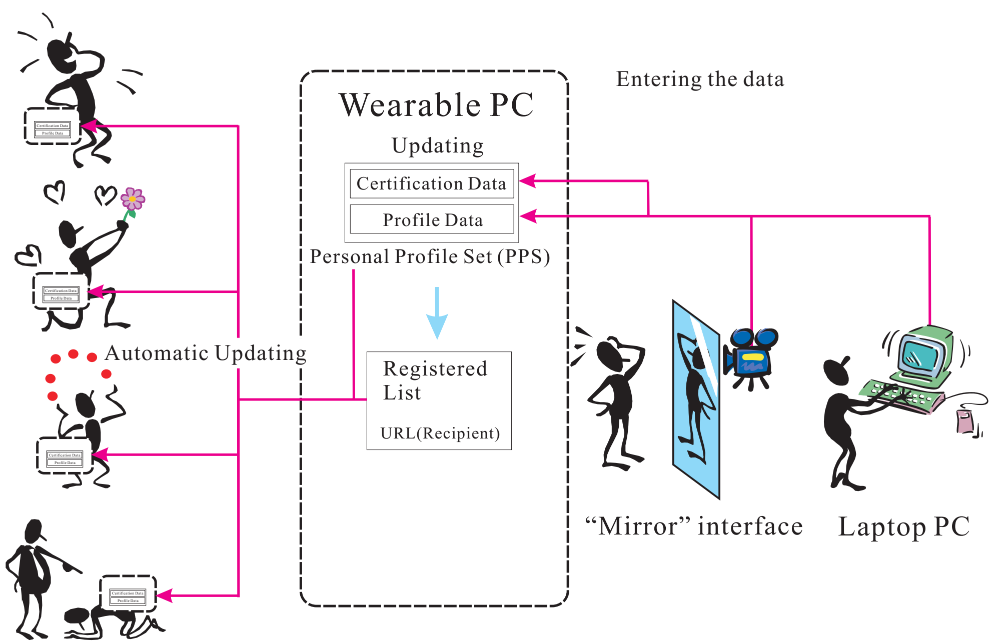


Figure 5.13. Automatic Update of a PPS

Figure 5.13 illustrates a general update method of PPS. This method uses a real-world interface method. The Nice2CU system automatically sends the entire group of registered recipients the target person's PPS when the certification data or the profile data is updated. This method provides the system with an direct referring operation to identify the target person who stands in front of the wearer. For example, the system recognizes the target person who meets after a long separation using the latest cartification data that consists of image data set for recognizing his/her face.

5.2.4 Rating Human Relations

In this thesis, I propose a rating method for clarifying human relations to recommend an augmented video memory, which will enable a user to recall information he/she wants to know, using human profile data and meeting logs. I termed this set of data and logs a "Dynamic Profile Network." This work is accomplished on a wearable system for a computational augmentation of human memory. A wearer views a video of his/her interest while talking with a target person who is standing in front of him/her. Viewing this video is helpful to the person when he/she cannot recall what topic is good for the context of a conversation. Also, selecting an appropriate video from a huge video data set recorded previously is difficult for the wearer. I categorize the wearer's friends into three types of friends. Additionally, we propose rating rules for evaluating human relations, and divide the rating rules into social rules and personal rules.

In order to recommend an augmented video memory to a user, I must implement human relations in the real world into the computational world. A node corresponds to a person and a link between two nodes denotes the relationship between corresponding persons. The goal of this paper is to clarify the structure of the Network and to construct a rating pattern on the Network. I first explain the structure of the Network among recipients: a target person, the recipients' friends, and the target's friends. Also, I discuss what kinds of topics the recipient and the target are interested in when they meet with other persons.

- *Recipient Friends*

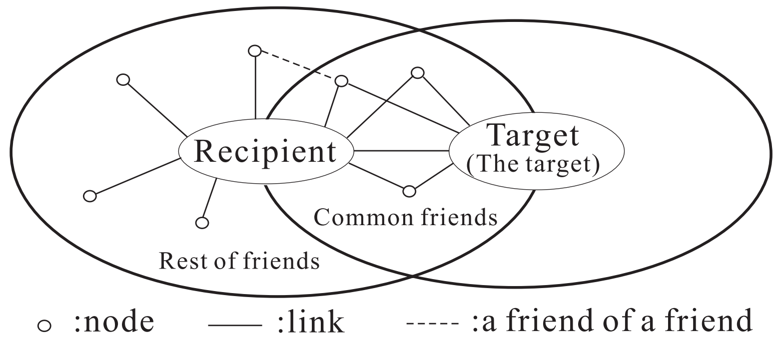


Figure 5.14. Set of Friends for the Recipient

• **Recipient's friends:** First, we consider a recipient's environment. We define a recipient's friends as the following three types of categories (Figure 5.14) from the point of view of topic selection in interpersonal communication.

The Target: The target represents a person who is standing in front of the recipient. The node that corresponds to the target has been directly linked to the node of the recipient. The relationship with the target is the first step toward better interpersonal communication. The main topic of their communication is to disclose his/her self.

Common Friends: Common friends between the recipient and the target are linked directly from both nodes. If common friends between the recipient and the target exist, then these friends belong to the same community. Many people enjoy talking about news surrounding a certain common friend. This news also corresponds to the news of a certain community.

The Rest of the Friends: The rest of the friends are the recipient's friends whom the target does not know. Most of the recipient's friends are supposed to belong to this category. Taking about a person in this category gives the target a chance to find a better friend.

- *Rating Rules*

I discuss the types of interests (rating points) in the "Dynamic Profile Network." Figure 5.15 depicts the kinds of rating rules in the Network. Personal rules include the recipient-side and target-side of personal rules. Here, "bias" in the figure represents the filter for deciding one's attitude to another person. Each type of rule, respectively, contains different intentions. I explain the characteristics as follows:

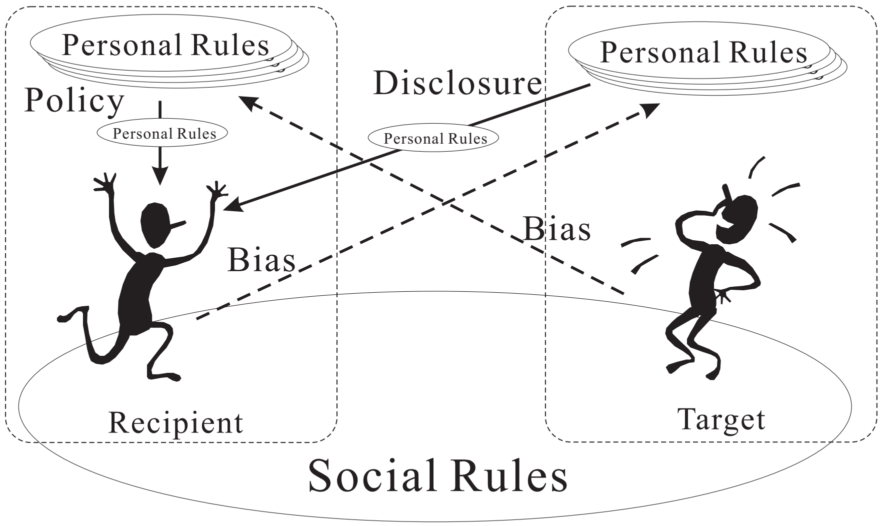


Figure 5.15. Two kinds of Rating Rules

• **Social rules:** The aim of social rules is for users of the *Nice2CU* to pick up information of interest from another person's profiles and meeting logs. For instance, if a person is to celebrate his/her friend's birthday, we first list example patterns of social rules. Social rules cover patterns of events. In each friend type, the system individually evaluates each pattern. Social rules are representatively composed of the following five types of interests:

S1 Basic Profiles: same birthday, birthplace, affiliation, hobby, address, etc.

- *The Target and Common Friend:* Suppose that the recipient and the target were born in the same town. If they are notified of this fact, they can chat about memories in their hometown. Suppose that both the target and a certain common friend are/were once colleagues or both were once students at the same school. If this is the case, then the recipient will have a good opportunity to talk about that topic with the target.

- *The Rest of the Friends:* Suppose that the certain rest of the friends have a similar hobby about which the target is also interested. The recipient can begin talking about the topic. Also, this may be a chance for the target to find a better friend.

S2 A Friend of A Friend

- *The Rest of the Friends:* A friend of a common friend between the recipient and the target would be someone who felt more kinship than the other the rest of the friends for the target by the target. Suppose that a certain rest of friends is a friend of a certain common friend and has the same hobby in which the target is interested. The target would want to meet with the friend more than with the other rest of friends who are not friends of any common friends.

S3 Messages

- *The Target:* Suppose that the recipient has borrowed a small amount of money from the target. If the recipient meets the target, the recipient can pay back the debt to him/her immediately by viewing the video message.

S4 Unbalanced Meeting

- *Common Friends:* Suppose that the recipient meets with a certain common friend more frequently than the target does. The target can request that friend to send a message such as: "I miss you. Please, call me," to the recipient.

S5 Temporal Events

- *The Target and Common Friends:* Suppose that a certain friend's birthday will be in a few days. The recipient and the target could plan to celebrate this certain friend's birthday. Also, suppose that the winter vacation is approaching and, the recipient, the target, and some of the common friends enjoy the same hobby of winter sports, e.g., skiing and snowboarding, then they can plan to go to a winter sports resort together.

• **Personal rules:** The recipients can also adopt their own attitude to the person they are meeting. Social rules cannot represent these kinds of personal properties. Such social rules are composed of the recipient-side of personal rules and the target-side of personal rules. The recipient-side rules are used to modify a recipient policy to the target. For instance, the recipient changes his/her own attitudes when he/she meets with his/her parents, teachers, or friends. The target-side of personal rules represent the desire of what topics he/she wants to be talked about by the recipient. For example, the target can set the following personal rule, "If a certain rest of friend is a different sex from the target and is younger/elder than him/her, then the rating of videos recording the friend should be improved by 100 points," if he/she wants to get a new girl/boy friend. Most of personal rules have complementary and adaptive roles. However, personal rules additionally have another important role. A person sometimes wants to have a completely different type of meeting with someone. The following "Anti-patterning" section provides several types of possible meetings.

Anti-Patterning: Anti-patterning represents the reverse of previously defined rules. For instance, "If a certain recipient's friend has different hobbies, then the rating of videos recording the friend should be improved by 15 points" is an example of anti-patterned rule when social rules include "If a certain rest of friends and the target have the same hobby, then the rating of videos recording the friend should be improved by 15 points" as one of rules. The *Nice2CU* system also computes an entire inversive evaluation in a certain anti-patterned rule. This anti-patterning allows for a high complexity of human networks. A complex human network leads to changes in dramatically clustered communities.

- *Model Policy*

• **Properties for Accessing Data:** In this paper, the *Nice2CU* system is allowed to rate human relations using the entire data of the target, the rest of the friends, the recipient's common friends, and the target's common friends. The data of the target's common friends are sent to the recipient's system. However, the data of the target's rest of friends are not allowed to be sent to the recipient's system because of privacy.

• **Rating Function:** Rating scores are computed in equation (5.1) in this paper. Note that, x_i is the point on the i-th rule, and w_i is a weighted factor on the i-th rule. N is a number of the rules.

$$Score = \sqrt{\frac{\sum_{i=0}^{n} w_i x_i^2}{N}} \tag{5.1}$$

5.3 Experiments

We have conducted a verification of the proposed idea using a Tabletop Role-Playing Game (TRPG) style experiment. We have employed a simulator imitating the world where all wearers use the *Nice2CU* system. We employed the simulator so that we can fundamentally correspond the verification of the rating method in the virtual world to verification in the real world, without an evaluation of user interfaces.

5.3.1 Methods

Three test subjects (players) participated in the experiment. Twelve characters were prepared, and the human network depicted in Figure 5.16 was prepared in

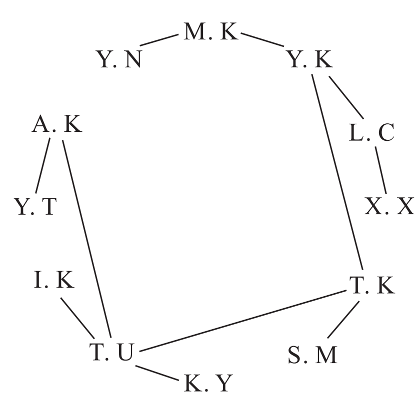


Figure 5.16. Initial Human Networks

advance. Although real profiles are employed in the network, human relations are assumed virtually by the experimenter. All nodes in the figure are set character's profiles. Relationships of a certain kind between two characters (node) are represented as a link in the figure. Each player acted as four characters. The period of the TRPG was eight-months virtually. The simulator can control all meeting of the events, and can select a recipient person and a target person. The simulator can register a target person as a new recipient's friend when he/she meets virtually with the target. In the experiment, the simulator recorded all the character's actions.

5.3.2 Results

Table 5.2. Pattern of Video Replay Times

Times	0	1	2	3	4	5	6
Amount	107	122	51	14	9	1	1

In this simulation, the simulator shows scenes 305 times to the players as a total number of recommendations. Table 5.2 shows the number of times each video was replayed. 35% of recorded videos were not replayed in the simulation. Characters totally left 20 messages to other characters, and 17 of 20 messages were replayed. There were 15 first meetings. The average number of replayed videos of a character was 25.67, and the standard deviation was 13.52.

Figure 5.17 shows the network at the end of the simulation. The average number of links in a node was 4.33, and the standard deviation was 1.56. Note that, the initial average number of links in a node was 1.83, and that the standard deviation was 1.03.

Table 5.3 shows the average of the replayed ranking ratio of recommended and replayed videos in the recommended top 20 videos. 49% of the top 1-5 ranked videos were selected in the simulation. 30% of the top 11-20 ranked videos were selected. The results show that the quality of the recommendations becomes unstable.

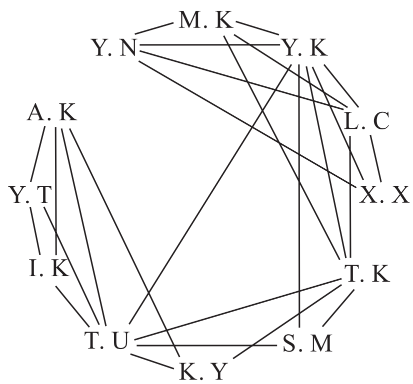


Figure 5.17. Simulated Human Networks

Table 5.3. Ranking Ratio

Rank	L.C	Y.K	T.K	Y.N	X.X	T.U	Y.T	I.K	S.M	A.K	K.Y	M.K	Avg
1-5	0.69	0.61	0.56	0.53	0.50	0.49	0.43	0.40	0.40	0.40	0.38	0.36	0.49
6-10	0.19	0.24	0.19	0.33	0.29	0.11	0.36	0.35	0.25	0.03	0.19	0.41	0.21
11-15	0.13	0.15	0.10	0.13	0.21	0.16	0.21	0.20	0.15	0.30	0.24	0.14	0.17
16-20	0.00	0.00	0.15	0.00	0.00	0.24	0.00	0.05	0.20	0.27	0.19	0.09	0.13

Table 5.4 shows the number of selected videos in each rank. We also categorize each meeting event in the simulation into the following three types: (1) Meeting, (2) Paperwork, and (3) General. To further investigate the simulation, we discriminate a general category among “Dinner,” “Sports,” “Birthday,” “News,” “Trips,” and “Greetings.” Table 5.5 shows the averages (M) and the standard deviations (SD) of the rank of the replayed videos.

Figure 5.6 shows replayed 17 message videos in the 20 left messages. All players chose a message when it showed in a cartain ranking at first time.

Table 5.4. Selection Patterns of Video Rankings

Ranking		1-5	6-10	11-15	16-20
Meeting		82	48	38	31
	(ratio)	0.41	0.24	0.19	0.16
Paperwork		12	5	8	3
	(ratio)	0.43	0.18	0.29	0.11
General	Toral	60	15	7	7
	(ratio)	0.67	0.17	0.08	0.08
	Amusement	5	1	3	1
	Birthday	7	0	0	0
	Dinner	21	5	1	4
	Greeting	3	1	0	1
	News	7	1	0	0
	Sports	13	4	2	1
	Trips	4	3	1	0

Table 5.5. M and SD of Rank of Replayed Videos

		M	SD
Meeting		8.15	6.87
Paperwork		8.14	6.19
General	Total	5.07	7.15
	Amusement	7.30	7.60
	Birthday	2.86	1.36
	Dinner	5.45	7.71
	Greeting	7.00	7.71
	News	2.38	2.98
	Sports	4.60	8.58
	Trips	5.38	3.06

Table 5.6. Ranking Patterns of Message

Rank		1	2	3	4
Meeting		10	0	0	1
General	Amusement	2	1	0	0
	Dinner	1	0	0	0
	Trips	1	0	0	0

5.3.3 Discussions

The most useful social rule is the S3(Messages) condition. In Table 5.6, 88.2% of S3 videos were selected when the videos showed the 1st rank. The message was used when a character had been on either a business trip or a sightseeing trip in the simulation. In the S5(Temporal Events) condition, for instance, a certain friend's birthday will be in a few days, and all videos showed within 1-5 ranks. This condition also gave a character the opportunity for a current conversation with a friend, and for a discussion about a future plan to meet the friend another day. Table 5.4 and 5.5 also show the performance of the model. However, Birthday and News in Table 5.5 represent low M and SD (Birthday M:2.86, SD:1.36; News M:2.38, SD:2.98). In contrast, Meeting and Paperwork showed a high ratio of the 16-20 ranks in Table 5.4, and M and SD (Meeting M:8.15, SD:6.87; Paperwork M:8.14, SD:6.19) were also high in Table 5.5. We believe that the model has the merit of recommendating video containing temporal contexts as in the S5 condition.

This work has the shortcoming of not providing an opportunity to find better friends using the *Nice2CU*, as shown in Figure 5.16 and 5.17. Human networks are constructed under a rule termed "scale-free networks" (Barabási, A.L., 2002) in our "small-world" (Watts, D.J. & Strgatz, S.H., 1998). The scale-free network, however, does not provide balanced opportunities among people who are referred to as the so-called "rich who just get richer." In terms of our model, it means a "rech person" represents a person who has more and more friends. In the simulation, we found that characters who had fewer friends had less chance to connect with other friends (Figure 5.17, initial SD: 1.03, final SD: 1.56) It is important for the characters to have an equal opportunity to make a friend even if the quality of information for their everyday lives would be low for the "rich." We believe that this service finally gives the user a good quality of information.

5.4 Related Works

By annotating a target person's profiles while the wearer meets the target, the wearer acquires three types of advantages: 1) recollecting the target's profiles, 2) creating a conversation, and 3) finding candidates of the target's new friend. These aims have been studied in wearable computing (Farringdon, J. & Oni, Y., 2000; Kato, T., et al., 2002a; Kato, T., et al., 2002b). However, these previous studies do not consider how the elements of human relations, i.e., personal information, should be managed. Hamasaki and Takeda also study personal networks on the WWW to find better friends (Hamasaki, M. & Takeda, H., 2003). Choudhury, T. & Pentland, A. (2003) investigate human networks in the real world using the sociometer. My study attempts to recommend video for finding better friends using the "Dynamic Profile Network" in the real world.

5.5 Summary

In this thesis I proposed the Card and Mirror prototype system of the *Nice2CU*. I discussed the easy registration method and the automatic update method. I also proposed a human relations-based rating method to recommend augmented video memory using a user's profile and meeting logs. I first structuralized the "Dynamic Profile Networks." I then defined two types of rating rules and proved the characteristics. Also, I showed the rating policy and the rating functions. Finally, I conducted the TRPG style simulation to prove the proposed rating model.

An additional issue to be addressed in my work regarding the *Nice2CU* is to provide a stabilized system for rating human relations in the real world. This issue, when settled, would be employed to investigate the usability of a video recommendation interface.

In this study, the *Nice2CU* only detects a target person from images captured by a wearable camera. 1) by detecting a person using his/her front face skin color, and 2) identifying him/her using a color information of the region under the face. The former one (the system can detect “a” person at a time) is not serious problem in this study because the user tends to pay attention to get the target’s information. The latter one is more serious. This study has not employed the latter problem as a main topic. I am, however, sure that this problem is a kind of inevitable topics for achieving the *Nice2CU* as an actual augmented memory system in the real world.

The topic of rating human relation has the other shortcoming. The problem shows low scalability of rating rules. The current *Nice2CU* module in this thesis has directly described both social rules and personal rules for conducting the TRPG style simulation. Nobody can write down his/her own personal rules and can get any new social rules. The problem shows that the *Nice2CU* must have a compiler/interpreter for editing social and personal rules.

Also, the *Nice2CU* gives the user an augmented memory when he/she meets the target. We can easily suppose that a lot of recorded augmented memories related to each target are stored in a wearable client computer. This fact means that it is difficult to retrieve desired augmented memories for the user. The *Nice2CU* must have an ability of selecting augmented memories to which the user wants to refer although the *Nice2CU* recommend augmtened memories he/she would be interested in by using the above social/personal rules.

Chapter 6

Ubiquitous Memories - Physical Object -

Main issues of this chapter are to establish a useful module to memorize a user's experiences using natual operations. In order to realize these operations, this chapter discusses a conceptual design first and then implement the module by investigating a mechanism of making an association between his/her experience and a real world object. The following topics are important to discuss memorizing, storing, recalling his/her experiences as an augmented memory on a wearable computer.

- designing a conceptual mechanism for facilitating arrangement of augmented memories in the real world
- implementing the designed module using real world interfaces
- investigating a performance of cognitive workload by using the designed and implemted concept

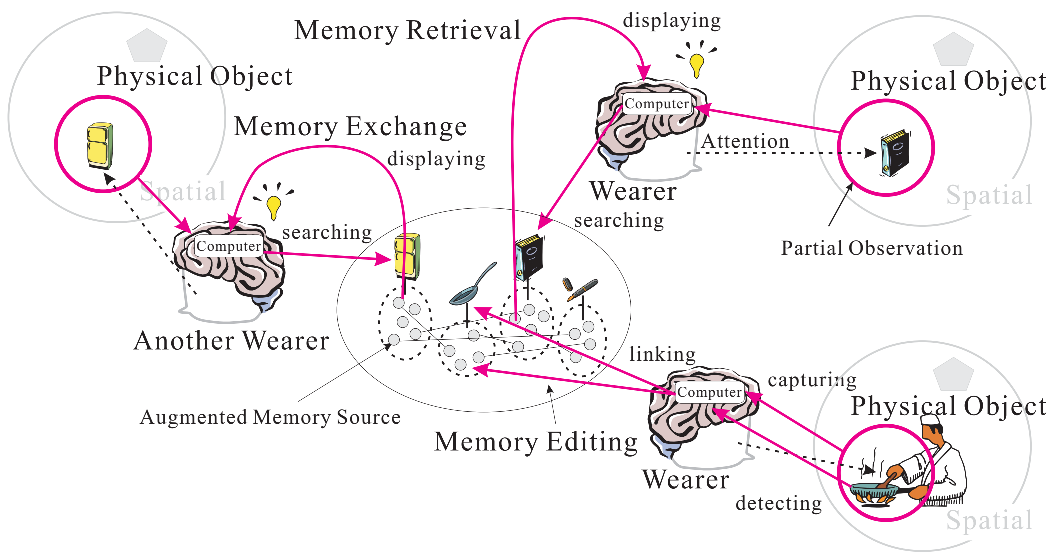


Figure 6.1. Object-triggered Augmented Memory Retrieval

6.1 Design

This chapter represents an example of the operational category of "Memory Retrieval," "Memory Exchange" and "Memory Editing" element mentioned in the chapter 3. Also, this chapter shows an example of the associable category of the "Physical Object" in the chapter 3. Figure 6.1 describes a model of a object-triggered memory retrieval. A user pays an attention to a certain object. The user tends to recall events associated with the object. A location-triggered augmented memory module detects userts intention and identifies the object, searches linked augmented memory sources with the object, and then the module shows the retrieved augmented memory source to the user. The user augments his/her memory by recalling an event the/she want to recall or he/she forgets.

This section introduces the *Ubiquitous Memories* modulesystem to support Memory Retrieval, Memory Exchange, and Memory Editing elements in the *SARA* (Kawamura, T., et al., 2002b; Kono, Y., et al., 2003). The primary motivation of the study is to enable wearers to manage everyday memories in the real world. In order to accomplish this motivation, I have proposed the concept of Ubiquitous Memories and developed a prototype system that can associate augmented memories with a physical object in the real world using a "touching" operation (Kawamura, T., et al., 2003a).

I propose a conceptual design to ideally and naturally correspond augmented memory to human memory. Conventionally, a person often perceives and understands a new event occurring in the real world by referring to his/her experiences and knowledge, and then by storing the memory of the event into his/her brain. He/she then obtains a novel and natural action for the event by analogically and metaphorically associating the event with previously occurring events. I believe that the acquisition of natural actions is important for realizing augmented memory. This acquisition positively establishes a "conceptual design" for seamless integration between human memory and augmented memory. In addition, the "Hand" interface has the potential for integrating augmented memory into objects.

Below I introduce the conceptual design of the *Ubiquitous Memories* system. The following procedures illustrate the conceptual design:

1. A person perceives an event via his/her body.

2. The perceived event is stored into his/her brain as a memory.

3. The human body is used as media for propagating memories, i.e., the memory travels all over his/her body like electricity, and the memory runs out of his/her hands. (Imitating this feeling, he/she can transfer the memory from his/her body to a physical object by "touching").

4. The transferred memory remains in the object.

5. He/she transfers the memory from the object to his/her body when he/she is interested in the object and touches it again.

6. Finally, he/she can recall the event.

6.1.1 Conceptual Design

In this chapter I define "Context" as information the human can sense in the real world, e.g., atmosphere, the observer's emotional condition, and the biometric

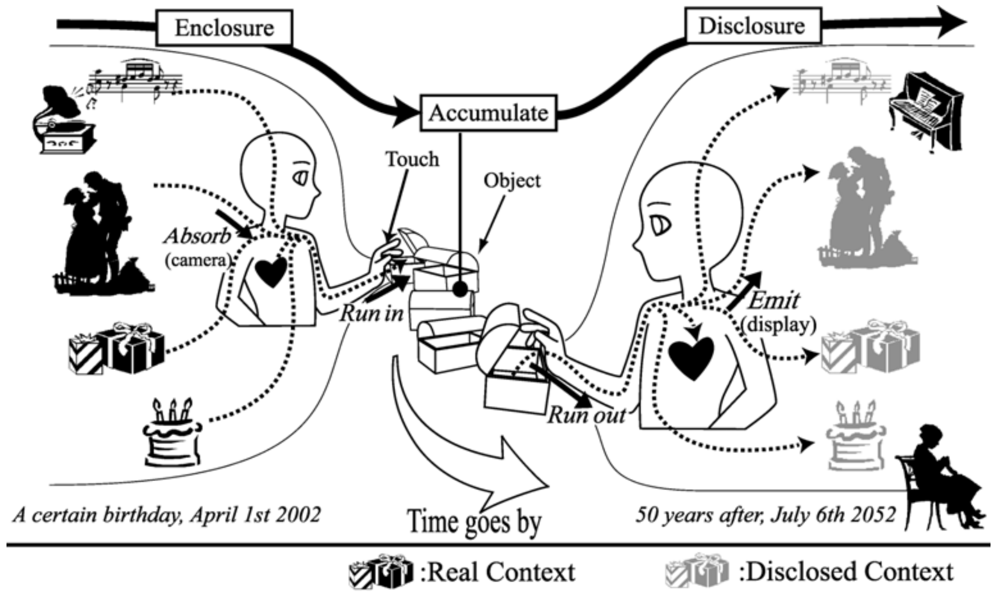


Figure 6.2. Concept of Ubiquitous Memories

states of the observer. Note that a context is not data like a video memory. The "Human Body" and an "Object" are important for my concept in realizing ubiquitous memories. Both the "Human Body" and "Real World Objects" are essential device/media for augmenting human memory in Ubiquitous Memories, i.e., the human body behaves as a device/media that associates video memory with objects. The terms of the conceptual actions shown in Figure 6.2 are defined as follows:

Enclosure action is shown by two steps of behavior: 1) A person implicitly/explicitly gathers current context through his/her own body and 2) He/she then arranges contexts as ubiquitous augmented memory with a real world object using a touching operation. The latter step is functionally similar to an operation that records video data to a conventional storage media, e.g., videotape, a CD-R/W, or a flash memory. The two steps mentioned above are more exactly defined as the following actions:

Absorb: A person's body acquires contexts from an environment, his/her own body, and his/her mind, as moisture penetrates into one's skin. Such an operation is called "Absorb" and is realized by employing real world sensing devices, e.g., a camera, a microphone, and a thermometer.

Run in: When a person touches a real world object, an augmented memory flows out from his/her hand and runs into the object. A "Run in" functionally associates an augmented memory with an object. In order to actualize this action, the system must recognize a contact between a person's hand and the object, and must identify the object.

Accumulation denotes a situation in which video memories are enclosed in an object. Functionally, this situation represents how the augmented memories are stored in storages with links to the object.

Disclosure action is a reproduction method where a person recalls the context enclosed in an object. "Disclosure" has a meaning similar to that of replay (for example, the way a DVD player runs a movie). This action is composed of the following actions: "Run out" and "Emit."

Run out: In contrast to "Run in," video memory runs out from an object and travels into a person's body. Computationally, the "Run out" 1) identifies the storage space where the augmented memories' linked objects are stored, and 2) these memories are retrieved from the Internet to a user's wearable PC. In order to achieve this action, the system needs contact and object identification functions such as "Run in." In addition, the system must have a retrieval function to refer to augmented memories associated with an object.

Emit: A wearer can restore contexts in an environment to his/her body, and mind. The system should employ devices, e.g., a video display and a headset, that can play back an augmented memory.

Enclosing an augmented memory in an object memory seeking behavior directly corresponds to an object-searching behavior where the object is associated with the memory in some scene. This correspondence gives a wearer more intuitive power to seek for an augmented memory using the principle of human memory encoding (Tulving, E. & Thomson, D.M., 1973). For example, suppose that a person won first prize in the 100-meter dash at an athletic festival and then got a plaque. A person can easily recall the event when he/she simply looks at the plaque because he/she has associated the event with the plaque in his/her mind. This associative ability is called the Encoding Specificity Principle. Two detailed characteristic traits exist for the principle when expressed in an object seeking action. One characteristic trait is the ability to recall an event or feeling or emotion by simply looking at or thinking about an object. This associative trait allows one to decide quickly what object he/she should seek. Another trait is the remembrance of placed location of the object. This trait allows a person to remember where he/she placed an object. These associative traits illustrate how we can easily recall an event by seeking out an object related to that event.

A person's touching operation is employed not only for realizing metaphors that a human hand implies ("Run in" and "Run out"), but also for naturally controlling an augmented memory system. Nonetheless, explicitly selecting an object for both human and computational devices makes it easy to for the user to express his/her distinct intentions by "touching."

6.2 Implementation

6.2.1 System

- - - - - - - - - - - - - - *Wearable Equipment for the Ubiquitous Memories system*

Figure 6.3 shows the equipment worn with Ubiquitous Memories. The user wears a Head-mounted Display (HMD; SHIMADZU, DataGlass2) to view video memories and a wearable camera (KURODA OPTRONICS, CCN-2712YS) to capture video memory of his/her viewpoint. The user also wears a Radio Frequency Identification (RFID; OMRON, Type-V720) tag reader/writer on his/her wrist. Additionally, the wearer uses a VAIO jog remote controller (SONY, PCGA-JRH1). To control the system, the wearer attaches RFID operation tags to the opposite side of wrist from the RFID tag reader/writer. The wearer carries a wearable computer on his/her hip. The RFID device can immediately read an

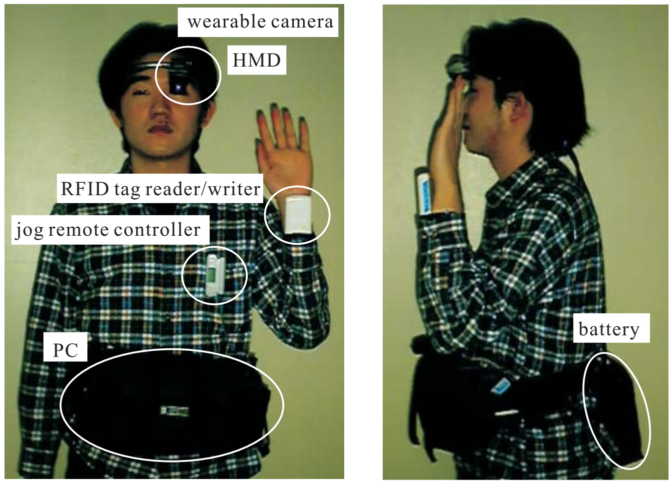


Figure 6.3. The Ubiquitous Memories Equipment

Figure 6.4. An RFID Tag Attached to a Cup

RFID tag data when the device comes close to the tag. The entire system connects to the World Wide Web via a wireless LAN.

\- *Real World Object Attached to an RFID tag*

I currently assume that an RFID tag is attached to/implanted in each real world object. Figure 6.4 depicts a sample image of an RFID tag attached to a cup. The Ubiquitous Memories system reads information from an IC of the RFID tag. Table 6.1 illustrates a tag data protocol to manage augmented memory. There are two facets of the RFID tag. One facet concerns identifying a certain object by attaching an RFID tag. I have employed a Serial Number (SRN), which is unique to each RFID tag before shipping, as an object identification number.

Table 6.1. An RFID Tag Information

Purpose	Object Identification Number	Data
Address for strage	SRN	URL
System Control	SRN	Operation Code

Another facet of the RFID tag is the data needed to: 1) address a server URL for storing and retrieving video memory by touching a real world object, and 2) send a command to the system by touching one of operation tags.

- *A System Configuration Diagram*

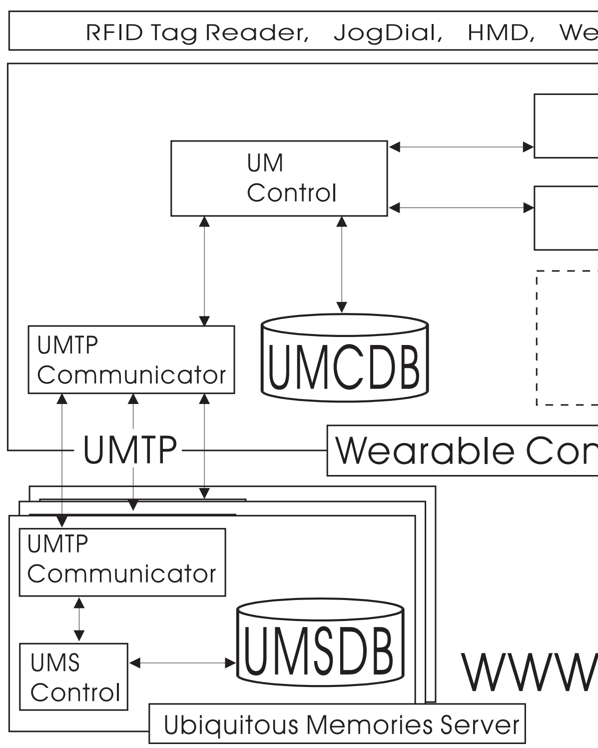


Figure 6.5. Ubiquitous Memories System Configuration

Figure 6.5 shows the system configuration of *Ubiquitous Memories*. This system is composed of a client system and multiple servers for a wearer. In this section I have termed the server a Ubiquitous Memories Server (UMS). The wearable computer plays the role of a client system. The core process is UM Control on the client system. The Video Buffer Control temporally manages a video memory before it is stored in a database. The Ring Buffer enables the wearer to choose two types of an enclosure operation. As a basic strategy, the wearer encloses a 10 second length video memory from the moment he/she wants to enclose it. On the other hand, the wearer can enclose a 10 second length video memory to a time when he/she wants to enclose it. This system has two types of databases. One is a privacy policy-oriented database that is placed in a wearable computer. I have named the database the Ubiquitous Memories Client Database (UMCDB). Another database is a public/group policy-oriented database. This database represents a server, which is termed the Ubiquitous Memories Server Database (UMSDB), in order to exchange video memory with other wearers.

Table 6.2. Type of Message and Data Transportation

Identification Part	Message Type
Message Part	OID, UID, AT, GID, TIME, (Command)
Data Part	(Video Memory)/(List Data)

Table 6.3. An Example of Message and Data Information

Identification Part	DATA
Message Part	0B8BE72400000009, 1000,1,9001, 20030909101231, OENC
Data Part	Data.avi

I have employed a special protocol named the Ubiquitous Memories Transfer Protocol (UMTP) to transport a video memory between a client and a server. Table 6.2 illustrates the types of message and data used for UMTP. Table 6.3 shows an actual example of the message and data. In the Ubiquitous Memories system, three types of data are transported: Message, Video Memory, and List

Data. In Table 6.2, a parenthesis shows that (Command), (Video Memory), or (List Data) are not needed in all processing. In the Message Type, the system can distinguish three states: "Only Message Information," "Attaching Video Memory," and "Attaching List Data." OID means an object identification number (SRN) recorded in an RFID tag. The UID is a user identification number. The AT shows an attribute of publication to other users. This attribute of publication is equal to a permission used to limit a user who is able to enclose video memory. The GID is a group identification number used for sharing video memories with friends, families, and co-workers. TIME is the time when the wearer first encloses a video memory. Command is equal to an operation code registered in an RFID tag. List Data is written down information (UID, AT, GID, TIME, and a filename of the video memory) of video memories that are enclosed to the same objects (OID).

6.2.2 Operations using Operating Tags

The Ubiquitous Memories system has six operational modes: ENCLOSURE, DISCLOSURE, DELETE, MOVE, COPY, and NONE. Note that the NONE mode means that the system reacts to a wearer's actions only when one of operation tags is touched. There are two basic operation tags and three additional operation tags for changing the mode. The wearer can select one of the following types:

- ENCLOSURE: By touching the "Enclosure" tag and an object sequentially, the wearer enclose video memory to an object. In this mode, the functions of "Absorb" and "Run in" are sequentially operated.

- DISCLOSURE: The wearer can disclose a video memory from a certain real world object. In this mode, the "Run out" function and the "Emit" function are sequentially operated.

Using additional operation tags, the wearer can treat a video memory in the real world like paper documents or data on a PC by using the following types of tag:

- DELETE: The wearer can delete a video memory enclosed in a certain object in the "DELETE" mode. This mode is used when he/she accidentally encloses an incorrect video memory, or when he/she thinks that a certain video memory is not needed anymore. (First, a video memory is run out from an object. He/she then emits contexts to the real world. Lastly, the video memory passes out of the object.)

- MOVE: This mode is useful when the wearer wants to move a video memory from a certain object to another object. For example, the wearer encloses a video memory to a notebook in advance when he/she is on a business trip. He/she rearranges video memories to each appropriate object after he/she comes back to his/her office. (First, a video memory is run out from an object. He/she then emits contexts to the real world. Lastly, he/she runs in the video memory to another object.)

- COPY: In this mode the wearer can copy a video memory to other objects. An event often has contextual relations with plural real world objects. This mode enables he/she to disperse a video memory to appropriate objects. (First, a video memory is run out from an object . The same video memory, however, remains in the object. He/she then emits context to the real world. Lastly, he/she runs in the video memory to another object.)

6.2.3 Operations for Selecting Publication/Reference Level

A wearer is allowed two ways to use a jog remote controller. One way is to set permission for publication by referring video memories that are enclosed to an object. Another way to use the job remote controller is to seek for an appropriate video memory from retrieved candidates.

A wearer must set a publication level attribute to a video memory to limit the people who can refer to the video memory when the user encloses it to an object. This publication attribute is particularly important when a wearer encloses a highly private video memory. Additionally, the wearer can select the reference level indicating the type of candidate video memories retrieved in a disclosure process. I have defined the following attributes:

- *Publication Level*

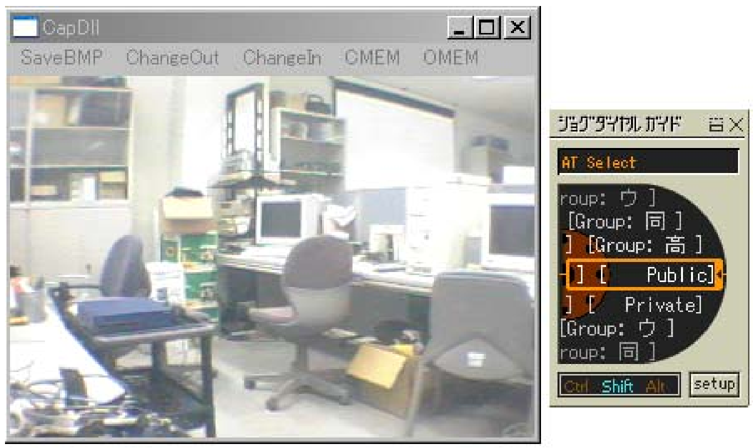


Figure 6.6. Selecting Publication Level

A publication level is set when the wearer encloses a video memory to an object. Figure 6.6 shows an example; he/she selects the attribute “Public.” Three types of publication levels exist:

- Private: Only the wearer who enclosed the video memory can disclose it.
- Group: Wearers who belong to a certain group can disclose it.
- Public: All users can disclose it.

- *Reference Level*

This level is selected when a wearer discloses a video memory from an object. In order to retrieve a desired video memory, he/she can reduce the video memories into disclosure candidates that are set at a certain publication level. Three types of reference levels exist:

- Personal: The wearer can disclose his/her own video memories.

- Group: The wearer can disclose his/her group's video memories published by the "Group."
- Public: The wearer can disclose all video memories that he/she is allowed to refer to by permission of the owners.

- *Finding an Appropriate Video Memory*

Figure 6.7. Selecting the Required Video

In the disclosure process a wearer can easily find the desired video memory if the number of memory candidates were enclosed in the touched object using the jog controller. As the number of enclosed memories increases, however, it becomes more and more difficult for the wearer to find the memory to be disclosed among the candidates even if a wearer is limited by selecting one of the reference levels. Figure 6.7 depicts how to find a video memory. In the example, the wearer views a snapshot of a video memory, which is activated by the controller, in the HMD. The wearer can change an activated snapshot when he/she turns the dial of the controller around. The wearer finally discloses a video memory by pushing down the controller when he/she finds a video memory of interest.

6.3 Experiments

I conducted two experiments to evaluate performances of the *Ubiquitous Memories* system (Kawamura, T., et al., 2002c; Kawamura, T., et al., 2003a). The experiments are to clarify a "succession of a principle" and a "comparison of mechanisms." The two experiments are detailed as follows:

- **A Succession of the Encoding Specificity Principle:** I must confirm the succession of the *Encoding Specificity Principle* in the *Ubiquitous Memories* system. Cognitive researchers have known that the principle is one of strong cognitive functions for memorizing and recalling extravagant contextual events in everyday life. The aim of the first experiment is to find out how much effect the system design has on maintaining contextual events. If the *Ubiquitous Memories* system succeeds the principle as an applicable augmented memory system, the results means that the system naturally has an effectiveness of the memory augmentation.

- **Recollection Efficiency of a Memorization Strategies:** I must compare the *Ubiquitous Memories* sytem with other memory externalization strategies. In terms of memory externalization, each strategy has a unique mechanism to memorize events and to recall events. I believe that investigating these mechanisms is one of important points to evaluate the *Ubiquitous Memories* system. The second experiment aims for finding what the *Ubiquitous Memories* system is different from other strategies and what memory strategy shows the most effective performance.

6.3.1 Influence of Context in Video to Human Memory

- *Methods*

This experiment was conducted at the Nara Institute of Science and Technology (NAIST) in Nara, Japan, among graduate students of the Information Science Department. 17 Japanese test subjects were participated in the experiment.

The experiment used pairs of an object image and a video in a trial (Figure 6.8). I set a notebook PC under laboratory conditions. A pair of an object image

Figure 6.8. Reference Display for Experiment 1

and a video was displayed on a PC monitor. The linked video is replayed automatically when the subject clicks the displayed image. The experiment contains 20 trials. In the 10 trials, each object in the displayed image is contained in the linked video. In the other 10 trials, each object and the video have no semantic relationship. In the experiment, each trial was alternately performed. Test subjects could watch a video 2 times in a trial, and then, the subject could proceed the next trial immediately.

After the subject finished all trials, there was a 3-minute delay period. In the period, the Japanese test subject had to read an scientific international journal paper and translate outloud each and every sentence from English into Japanese. The subject then answered a questionnaire. In addition, after a 1-week delay period, the test subjects had to answer the same questionnaire again. 11 questionnaires were returned.

The questionnaire contained 20 recall questions in Japanese. Each question showed an object image used in the experiment. The subjects filled in as many answers as they could in less than 10 minutes. It should be noted that the sequence of questions was different from the sequence of trials. The test subjects could answer these questionnaires in a random order.

- *Results*

This section analyzes the results of the 17 (3-minute delay) and 11 (1-week delay) questionnaires that I collected from the Japanese graduate students of the Information Science Department at the NAIST in Nara, Japan.

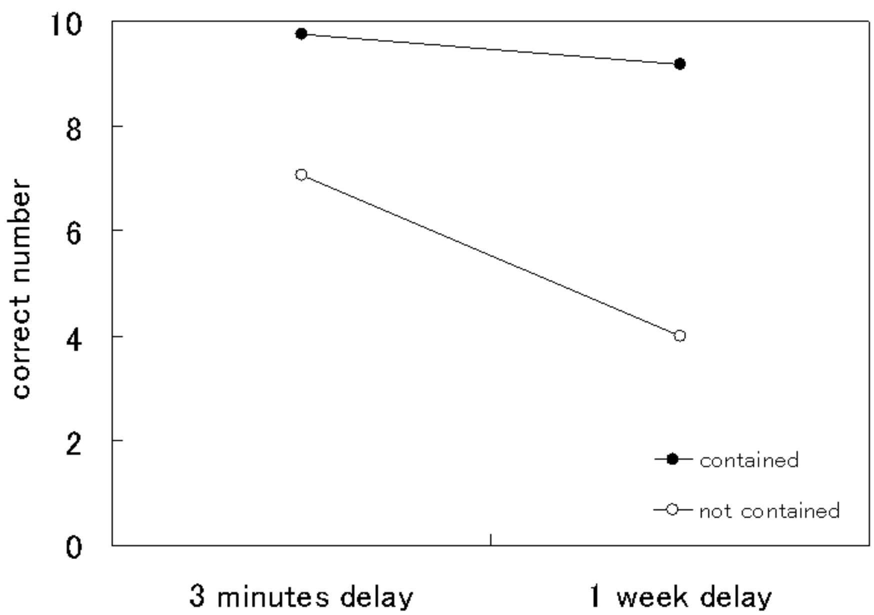


Figure 6.9. Context Effect

Figure 6.9 shows the average number of correct answers. The left side of the figure shows the average number of the correct answers in the 3-minute delay questionnaires. The right side illustrates the average number of correct answers in the 1-week delay questionnaires. Black dots result of the questions for images with videos containing images themselves. White dots illustrate for images, which were not contained in the video.

In the 3-minute delay questionnaires, the average number of black dots is approximately 1.4 times greater than the average number of white dots. This result shows a significant difference ($p<0.0005$). In the 1-week delay questionnaires, a difference between the average number of black dots and the average number of white dots widens to approximately 2.3 times. This result shows a significant difference ($p<0.0\,005$).

6.3.2 Recollection Efficiency of a Memorization Strategy

- *Methods*

This experiment was conducted at the NAIST (Nara Institute of Science and Technology) in Nara, Japan, among graduate students of the Information Science Department. 20 test subjects were included in this experiment.

For materials, I used 10 physical objects that had no contextual relation to each other. I also used 10 portraits of unfamiliar persons, and two sets of 10 playing cards composed of the numbers 1 through 10. I conducted the experiment under laboratory conditions. One experiment was composed of a memory test and a recall test. In the memory test, the subject memorized 10 trials. In the recall test, the subject answered a questionnaire.

In a trial of the memory test, the subject was first shown a pair consisting of an object and a portrait. The subject then selected one of the corners of the portrait. Finally, the subject was shown the predetermined pair of playing cards. The subject was allowed to look at these numbers within 30 seconds. The subject had to memorize the object and portrait pair, and the corner of the portrait and two card numbers as a real world experience including narrative contexts. The subject continuously tried to memorize all trials. All subjects had to do two experiments within the following four conditions:

C1: use only human memory (learn by heart)

C2: use only facial characteristics (record with a paper and a pen)

C3: refer to photo album type portraits that were used in the memory trial

C4: use the Ubiquitous Memories system to refer to portraits in the recall test

Test subjects were divided into four groups. Group 1 did two experiments using conditions C1 through C3. Group 2 experimented using conditions C3 through C1. Group 3 experimented two times using conditions C2 through C4. Group4 did two experiments in conditions C4 through C2. In the recall test, a questionnaire contained 10 recall questions. The subject was given one object image in each question. There were three empty boxes (portrait, corner and card numbers) in a question. The subject then selected a portrait ID from a list having 40 portraits, marked a corner (Left-Top, Left-Bottom, Right-Top, Right-Bottom), and wrote down two card numbers. The subject was then given a list of 10 portraits used in the memory test only in condition C3. All subjects filled in some or all answers within 10 minutes. The question sequence was changed from the trial sequence in the memory test. All subjects ware allowed to answer the questions in a random order.

- *Results*

The results were taken of the 20 questionnaires collected from the graduate students of the Information Science Department in NAIST. Table 7 illustrates recall rates from the 20 questionnaires. In this section I defined N, P, B, F and '. N, which is a percentage of errors, as follows: no answers regarding a portrait, a card and card numbers were correct answers on several questions. P shows that the answer regarding the portrait was correct. B shows that the answer of a corner was correct. F represents the answers of card numbers that were correct. X' (X is either P, B or F) represents the answer of a question X that was not correct.

In Table 6.4, N and P show a significant difference among the four test conditions (**: $p < 0.001$). P by C4 is, however, not 100% because of system error. In the sum of P'BF', P'B'F, and R'BF, I can see the influence on the difference among the test conditions (C1: 25.0%, C2: 14.0%, C3: 11.0%, C4: 4.0%, $p < 0.001$). The sum of PBF' and PBF shows the transparency in the different test conditions (C1: 36.0%, C2: 38.0%, C3: 49.0%, C4: 56.0%, $p < 0.1$).

Figure 6.10 shows an F-flow model to explain dependency among P, B, and F in Table 6.4. I can investigate the test subject's remembrance strategies for the experiment. I introduce O that represents an object as a question in the questionnaire. I envisioned a flow rate of a flow path for the model. A flow

Table 6.4. Recall rate

	C 1	C 2	C 3	C 4
N**	24.0%	31.0%	10.0%	2.0%
PB'F'	11.0%	8.0%	19.0%	19.0%
P'BF'	12.0%	9.0%	5.0%	3.0%
PBF'	23.0%	20.0%	32.0%	31.0%
P'B'F	8.0%	4.0%	3.0%	1.0%
PB'F	4.0%	9.0%	11.0%	19.0%
P'BF	5.0%	1.0%	3.0%	0.0%
PBF	13.0%	18.0%	17.0%	25.0%
P**	51.0%	55.0%	79.0%	94.0%
B	53.0%	48.0%	57.0%	59.0%
F	30.0%	32.0%	34.0%	45.0%

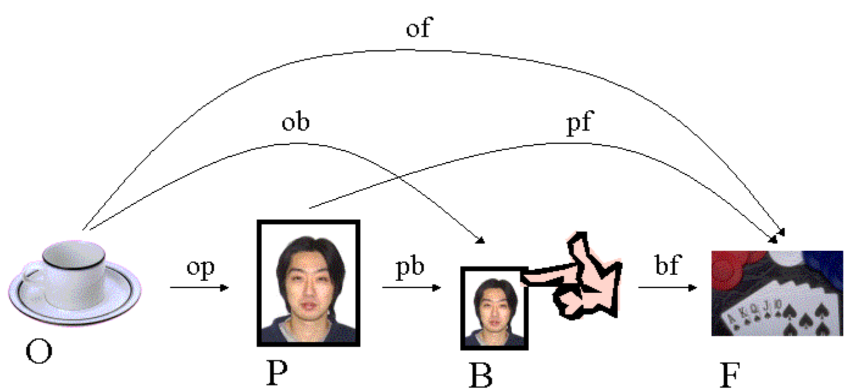


Figure 6.10. F-flow Model

Table 6.5. F-flow path ratios

	C 1	C 2	C 3	C 4
of	17.9%	9.5%	7.1%	2.5%
op**	74.7%	88.9%	89.5%	97.5%
ob	7.3%	1.6%	3.4%	0.0%
pf	44.7%	50.9%	40.5%	38.5%
pb	30.0%	38.0%	49.0%	59.0%
bf	37.3%	39.6%	52.4%	59.0%

pattern in the model corresponds with an associative pattern from O to F. The outflow of recall from O and the inflow to F are both 100% as a total. Table 6.5 shows the representation of the F flow in the experiment. The path of op, ob, of, pb, pf and bf show the percentages for the flow rates.

In an analysis of the variance to the result of the experiment, op shows significant differences (p<0.0005).

6.3.3 Discussions

I need to cofirm the succession of the *Encoding Specificity Principle*. Figure 6.9 depicts that the *Ubiquitous Memories* system allows users to effectively perform contextual events associated with real world objects. The result just means that the effect cannot perform when an event has no contextual relations to the objects.

I also need to investigate which kind of memory aid strategy performed best. Table 6.5 shows that the *Ubiquitous Memories* system was the most effective. The differences are especially clear in the result of N and P. In the result of PBF (C1: 13.0%, C2: 18.0%, C3: 17.0%, C4: 25.0%), C4 showed the relationship of an object associated with a portrait. Additionally, the sum of P'BF', P'B'F, and P'BF (C1: 25.0%, C2: 14.0%, C3: 11.0%, C4: 4.0%, $p < 0.001$) shows a good result for the *Ubiquitous Memories* system. Additionly, Table 6.6 describes that the *Ubiquitous Memories* system (C4) also represents the best performance in a recollection cost of externalized memory in pb and bf by reducing the cost for remembering op on the Table 6.5. The results of C4, however, includes accidental system errors. The errors show N, P'BF', and P'B'F in Table 6.5 and of in Table 6.6.

In the experiment, the system showed the following two considerable results:

1. The system effectively and clearly supports users' contextual event association with real world objects by using the *Encoding Specificity Principle*.

2. The result shows that "Enclosure" and "Disclosure" operations, which enable wearers to directly record/refer to a video memory into/from an object, is effective enough to make human memories ubiquitously spread in the real world.

I believe that the system is more useful than conventional externalized memory-aid strategies. Increasing the workload adds knowledge about how best to conduct oneself in a certain situation or events in our increasing complicated lives. The former result means that a user can make ubiquitous memories without special overloads using the proposed system.

6.4 Related Works

Rekimoto et al. proposed a system for browsing information that included the physical context of a user's current situation (Rekimoto, J., et al., 1998). Their aim is to develop interactive technologies in physical environments. However, neither a computational augmentation of human memory nor a development for operating a memory aid system was considered. Michael et al. proposed artefact-centerd computing, called MediaCups (Michael, B., et al., 2001). This system needs electric power to use digital artefacts.

6.5 Summary

I introduced the *Ubiquitous Memories* that enables wearers to make ubiquitously spread augmented memories by employing real world objects. The *Ubiquitous*

Memories is implemented by employing a natural operation "touching." The *Ubiquitous Memories* has been composed of a wearable camera, with an RFID tag reader worn by the user, and an RFID tag attached to each real world object. This system has enough basic operations to directly enclose/disclose a video memory into/from an object.

The shortcomings of this study include the necessity for recording a fixed length video data (10 seconds) as an augmented memory in a "Enclosure" process. Also, the *Ubiquitous Memories* cannot retrieve a proper augmented memory from a huge collection of enclosed augmented memories in an object. In the *Ubiquitous Memories*, the wearer selects an augmented memory from listed all memories linked with the object with a jog remote controller. I believe that the latter problem would be solved by recmmending candidate augmented memories for recalling a past event the wearer want to do that.

Furthermore, an object selection problem remains in both operations of "Enclosure" and "Disclosure." In the case of the object selection, a wearer might worry about selecting an object to enclose a video memory. In order to resolve these problems, I should structuralize contextual relations among persons, objects, and contexts (Michael, B., et al., 2001; Michael, B., et al., 2002), and investigate the similarity of enclosure and disclosure patterns among similar objects and among non-related objects.

Lastly, I believe that the "Memory Exchange" function will be an important issue in the near future because cooperative recall is superior to individual recall in the memory process (Takatori, K., 1980).

Part IV

Conclusions

Chapter 7

Conclusions

7.1 Concluding Remarks

The augmented memory research area has been standing in front of the entrance. I found that any research field (e.g., psychology, sociology, brain science, information science, and information technology) cannot accomplish computational augmentation of human memory by itself. I also found that extravagant problems lie ahead of each research field. In order to realize augmented memory as a key-technology/killer-application of both wearable computing and ubiquitous computing, I hope that the research area should be studied by many researchers from several research fields. I believe that the study of augmented memory is an excellent and applicable test bed for integrating information spaces in such fields as psychology, sociology, brain science, information science, and information technology, etc. into the complex real world. I also believe that the research area provides a good chance to collaborate the different research fields.

A considerable problem also remains as a general issue. Does "Augmented Memory" truly enrich our everyday life? Does the application of "Augmented Memory" reduce our cognitive workload of memory activity seriously? Actually, I definitely know that there are additionally any other kinds of negative anxieties and disbeliefs. However, all augmented memory researchers must discover "**cog-**

nitive tricks" that can effectively reallocate resources of memory activities. I believe that the following research procedures are the essencetials for augmented memory.

1. Investigation of memory activities
2. Discovery of effective computational augmentation (trick)
3. Formalization of discovered tricks
4. Design of a module using formalized tricks

7.2 Contributions

The contributions of this thesis are composed of the design of augmented memory and the implementation of some modules of the SARA framework for realizing augmented memory. Detailed contributions are shown as follows:

- *Design*

The contributions of the design for augmented memory are divided into two kinds of details. One is redefinition of augmented memory. Another is desgn of the framework of for the augmented memory system of SARA.

- **Redefinition of Augmented Memory:** I discussed the definition of Rhodes's augmented memory and clarified structures of the memory corresponded to human memory process. In the encoding process, I explained merits and lacks of processing augmented memories on the encoding, storage, and retrieval process. I explained the "Automatic" and "Semi-automatic" encoding process, and the "Subliminal" and "Cognitive" encoding process. In the storage process, I also introduced the "Personalize," and "Communitize," storage process. Lastly, I discussed the "Subliminal" and "Intuitive" retrieval process. The discussions make us to construct the concept of augmented memory and the after-mentioned framework of SARA.

- **Proposing the Framework of SARA:** I discussed what kinds of operatons and concepts the augmented memory must have. I proposed two kinds of categories. One is "Operational Category": Memory Retrieval, Memory Transportation, Memory Exchange, and Memory Editing. The categories provide us a direction of fundamental function design of augmented memory modules. Another is "Associable Category": Spatial, Physical, Temporal, Behavioral, and Psychal. The categories give us a direction of elemental structure design of augmented memory modules. I believe that the framework is a base concept that enables us to integrate among individually implemented augmented memory modules.

- **Design of Fundamental Modules on the SARA framework:** The contributions of the design shows three types of coceptual mechanisms based on the associable categories: 1) continuity and spatiality characteristics of events linked with physical space, 2) dynamics of person's information and human relations, and 3) physical property with memory between human body and real world objects.

  1. *Spatial Event Transition* The design of the module contributes to traceability of spatial event against continuous event transitions. The design enables the augmented memory module to retrieve similar scene in the real world, so that the module enables a wearer to know whether the recalled past event was truly occurred at the place where he/she is or it is just a certain deja vu phenomenon.

  2. *Dynamic Profile Network using Person's Information* The contribution of the module design is to simplify a structure of person's information and human relations. The success of the design is on the structuring human relation limited to set of friends and the formalizing rules of human relations divided into social rules and personal rules. The establishments of a chance to register a new friend, and to update person's information also achieve a getting facilitation of module operations.

3. *The Relationships among Human, Object, and Memory* The design of the module gives us a contribution of a conceptual design. The design is implanted a cognitive mechanism of "Encoding Specificity Principle." The module also employed other considered designs using cognitive characteristics, e.g., analogy, metaphor, and arrangement of memories using the relationships between a physical object and a space ("a toothbrush" and "a rest room"). The contribution shows that these cognitive characteristics enable us to minimize amount of applying computational techniques.

- *Implementation*

The contributions of the implementations for augmented memory are divided into three kinds of details. I representatively introduce the three contributions as follows:

- **Spacial - Residual Memory:** The module has four contributions: 1) using two mono-axis gyro sensors to exclude motion information for image matching, 2) also using the sensors for segmenting video scenes from a wearer's motion, 3) using a hybrid-space high-speed video scene tracking method for retrieving a video scene, and 4) using HSV probabilistic density histogram for image matching. In the work, I conducted four kinds of experiments in each technical contribution.

  1. *Image Matching using Gyro Sensors* The method resolves robust image matching including wearer's motion information and moving objects in an image. Conventional image matching methods have not respectably excluded both motion areas on the image. The image captured from a wearable camera includes three types of information: 1) static area information, 2) moving object area information, and 3) the wearer's head/body motion information. My method can respectively discriminate the combined information into the three types of information. Finally, the module can compute a similarity of static

area information (location information) between a current image and recorded image.

2. *Video Segmentation using Gyro Sensors* I employed the moving average method to segment video scenes by analyzing trends of a wearer's behavior. The method has been used in the field of the derivative engineering to predict a trend of varying stock market prices. After all, the method can successfully be adopted the analyzing of the wearer's behavior.

3. *Hybrid-space Video Tracking* The most important contribution of the method is to model a scene tracking-based video retrieval strategy. Characteristics of space are continuity and spatiality. The method is designed by using the characteristics. The mothod prepared two feature spaces. One is continuous feature space. Another is distributed feature space. The scene tracking prosess traversing between the two feature spaces allows the real-time video retrieval from huge video data set in the real world.

4. *Image Matching using Probabilistic Histogram* The method gives us to reduce influence of camera noise to recognize similar scenes. The method employed a probabilistic representation of histogram on HSV color space. The histogram on HSV color space is one of conventional feature to recognize similar objects. The method shows better performance than the conventional hue histogram compared with the same size of feature space althogh the computation time is increased.

- **Physical Human - Nice2CU:** The module has two contributions: 1) designing real world-oriented interface of augmented memory using a business card and a mirror, and implementing easy registration and automatic update methods in the real world, and 2) rating human relations using human profile data and meeting logs. I also conducted the Tabletop Role-Playing Game (TRPG) style simulation for proving the rating model.

  1. *The "Card and Mirror" Interface* The contribution of the interface is

to clarify how to put computational augmentation into a mechanism of interpersonal communication in the real world. Exchanging business cards is often seen in the Japanese business scene. The exchange is an opportunity to disclose and register the target person. Also, mirrors are often used for checking our own appearance before going out from home. Such a situation gives the user the opportunity to update the user's information. The two tools totally allow us to employ augmented memory to trigger our recollection of person's information.

2. *Rating Human Relation* The study contributes for a conversation in the real world and for finding a better friend. The definition of the recipient friends enables us to establish the framework of the social rules and the personal rules. Also, I represented the patterns of social rules. The study develops a scientific framework for learning the human network dynamically updating person's profiles and interacting each other in the real world.

- **Physical Object - Ubiquitous Memories:** The module has two contributions: 1) proposing the conceptual design and implementing the design using the RFID tag system to associate augmented memory with real world objects, and 2) evaluating the module with two experiments.

1. *Implemented module* The module equipment is composed of light weighted devices and a computer. The wearer can continuously use the equipment for several hours in everyday life. The RFID tag system employed for operating the module enables the wearer to associate augmented memory with a real world object by a "touching" operation using his/her hand attached to the RFID tag reader/writer. The wearer can associate augmented memories with a real world object using the latest version of the *Ubiquitous Memories*. Also, he/she can either copy, move, or delete associated augmented memories like as using physical paper documents, and digital documents. Additionally, the wearer can set a publication level to limit the people who can refer

to augmented memory, and can select a reference level to reduce candidates of augmented memories to which he/she want to refer. Lastly, the *Ubiquitous Memories* provide the Memory Exchange function to wearers in the real world, and they could refer to an object's history as a digital nostalgia.

2. *Experiments* The contribution of the two experiments gives us considerable results for accomplishing the augmented memory. One is to prove a succession of "Encoding Specificity Principle" into the *Ubiquitous Memories.* The results show that associating augmented memories with real world objects enables users to directly recall events involving in the object. Another experiment is to evaluate a performance of the *Ubiquitous Memories* in comparison with other memory externalizing strategies, e.g., a learning by heart, a memorandum, and a photo album. The results presents that the *Ubiquitous Memories* has better performance to manage memories associated with physical objects.

- **Implementation of Operational Category:** The implementations of this thesis show how to realize proposed "Operational" categories and base of "Associable" categories. In the Memory Retrieval, the *Residual Memory*, the *Nice2CU*, and the *Ubiquitous Memories* enable us to respectively recall a location-based memory, a human-based memory, and an object-based memory. Especially, the *Ubiquitous Memories* chapter describes a cognitive encoding process using the "Encoding Specificity Principle" to retrieve memoris. The result gives us a direction of non-computational or low-computational human memory augmentation for a memory retrieval process. Also, the *Ubiquitous Memories* could be partially integrated into the *Nice2CU* because of the similar module design employed in each module. The module integration expands our memory activity using augmented memory. In the Memory Exchange, the *Ubiquitous Memories* has the ability that the wearer can exchange augmented memory sources with an other wearers via augmented real world objects. In Memory Editing, the wearer can edit and update augmented memories (human profile data, "Profiles") using the "Mirror" interface of the *Nice2CU*. The *Nice2CU* also gives him/her another Memory Editing operation implicitly as "Experiences" when he/she meets a target person, and explicitly as "Messages." The editing operation makes the information relatd to human-human relations as "Human Relations." The *Ubiquitous Memories* enables the wearer to edit augmented memory by a linking operation between augmented memories and real world objects. The wearer exhaustively makes use of all edited augmented memories.

- **Implementation of Associable Category:** Spatial element, the Physical Human element, and the Physical Object element are implemented as fundamental elements of associable categories. Implementing the elements means that augmented memory modules can manage the static situation of the world as augmented memories. And then integration of implemented modules gives us operations of augmented memory representing the static world. However, the *Residual Memory* can retrieve previously recorded scene, and then the *Ubiquitous Memories* can associate the scene with a real world object. The sequential operation means that the wearer can use a dynamic operation of augmented memory. Additionally, the *Nice2CU* can also operate dynamic augmented memory using the "Dynamic Profile Network."

7.3 Further Directions

The research for the augmented memory system has grown in recent years. However, most augmented memory researchers have dedicated themselves to resolving problems similar to previous problems, and have analyzed simplified human traits of memory activities. Recently, such research has been directed primary toward clearing up issues related to wearable computing such as: investigating relationships between video data and brain wave data and supporting memory activity via a subliminal effect. Exchanging experiences is also a key-topic in realizing augmented memory systems.

7.3.1 SARA Framework

The future research direction of the SARA framework includes connection-ability problems among super-distributed augmented memory modules. One of the most pressing problems for consideration is the ultimate realization of the operation of "Association (surfing on the augmented memory)." In order to accomplish "Association," an augmented memory system would require standard notation rules to connect a huge amount of modules.

Another difficult issue is the exchange of augmented memories among various people. In order to exchange augmented memories, a system has to convert information from personalized augmented memories into a public domain for other desired users. Most studies for augmented memories, however, considered all human experience from homogeneous viewpoint. Each person has his/her unique perception even if the persons share an experience at the same time and in the same place. I believe that exchanging useful augmented memories to realize information conversion requires contributions of relations or associations between multimedia data management strategies and human memory traits.

7.3.2 Spatial - Residual Memory -

In this work, the module has three types of techniques. First, the module can match similar images captured from a wearable camera set at the center of wearer's face. The image includes his/her motion information. The module excludes his/her motion information, and detect similar landscape image using gyro sensors. Second, the module segments video scene using only gyro sensors. The experimental results mean that the module can employ a motion-oriented video segmentation method. Third, the module shows a video to the wearer using a high-speed video retrieval method. The contribution totally shows the wearer refer to an associable video data using only his/her gazing type of the video retrieval operation.

The contribution, however, also means that the module cannot show an appropriate video from a huge video data. The technical further direction of this work is an on-demand recommendation of videos in which the user would be interested from huge video data set. In the case, the module could be allowed parallel tracking of similar video scene. I believe that the resolution of this issue is not in sequential video search methods.

In addition to the technical directions, design issues remain in the work. The module enables the wearer to confirm whether a certain deja vu phenomenon is true or not. However, the module cannot differentiate video captured at a certain location from similar videos captured at different locations. The location-based augmented memory module must have absolute position-based video retrieval methods. The module should be investigated positive and negative affects by psychological experiments.

7.3.3 Physical Human - Nice2CU -

Main topic of the *Nice2CU* module is how to manage person's augmented memories in the real world. The module provides functions of Memory Retrieval, Memory Exchange, and Memory Editing. First, the module identifies a person who is standing in front of the recipient. Second, the module enables the wearer to easily register a target person who meets with the wearer in the first time, and to automatically update his/her profiles. Third, the wearer can use a video recommendation function.

I do not know what kinds of meeting data give the wearer a profitable day. Analyzing human relations are important issues to keep friends for a long time, and to get good opportunity to find better friends. The conditions make strong and stable social connections. In order to establish strong and stable social connections, it is important for us to support a exchange of person's information. Especially, I am investigating a method to measure closeness between persons in the real world. The module should clarify the human network, and then the module must track dynamically changed human relations day by day. For instance, I should investigate how to estimate closeness between the recipient and the target. Also, the module should define profile properties, and should have real world-oriented interface to enter the profile data. Additionally, the module has technological problems on the method how to identify a target person who is standing in front of the recipient. The following three main research topics are in preparation for summarized future works.

1. Genneralization: analyzing rating patterns
2. Satisfaction: improving rating functions
3. Facilitation: developing rating rule scripts

The additional obvious shortcoming of this study includes how to reject illegal access from an RFID business card. This privacy issue is one of important problems for realizing an applicable wearable system.

7.3.4 Physical Object - Ubiquitous Memories -

This work provides an effective design of memory arrangement strategy. The module can associate wearer's viewpoint videos with real world objects. The wearer can refer to videos from the objects, and also can select a publication/ reference level of an associated video for a social use of the *Ubiquitous Memories* module. I believe that the module has a fundamental mechanism for a real world-oriented interface using physical objects.

Further direction of this work is to accomplish a video recommendation method from huge video data set associated with each real world object. However, it is difficult for us to realize the method. First, I have not yet investigated what kinds of relations are between a certain characteristic real world object and associated video data. Additionally, I have not known whether I could discriminate the relations into personal relations and social relations. The characteristic differences of relations would give users object-specific video recommendation methods. I consider there are at least two methods to recommend video. One is to symbolize roles of an object and record interaction logs between a certain object and wearers like the relationships rating method of the *Nice2CU* module. Another is to analyzing correlations among associated videos with an object. A correlation among a certain video and other videos is also important for the issue. In order to find the above relations and correlations, I should investigate the mechanisms with several test participants and almost objects attached an RFID tag in the real world for a long time.

References

Abowd, G.D. and Mynatt, E.D. (2002). The Human Experience. *IEEE Pervasive Computing*, March, pp.48–57.

Aizawa, K., Ishijima, K. and Shina, M. (2001). Automatic Summarization of Wearable Video-Indexing Subjective Interest. *Proceedings of the 2nd IEEE Pacific-Rim Conference on Multimedia, Springer LNCS2195*, pp.16–23.

Aoki, H., Schiele, B. and Pentland, A. (1999). Realtime Personal Positioning System for Wearable Computers. *Proceedings of the 3rd International Symposium on Wearable Computers (ISWC'97)*, pp.37–43.

Barabási, A.L. (2002). Linked: The New Science of Net-works. *Perseus Publishing.*

Barclay, C.R. (1996). Autobiographical Remembering: Narrative Constraints on Objectified Selves. In D.C. Rubin (Ed.), *Remembering Our Past: Studies in Autobiographical Memory*, New York: Cambridge University Press, pp.94–125.

Barsalou, L.W. (1988). The Content and Organization of Autobiographical Memories. In U. Neisser & E. Winograd (Eds.), *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*, New York: Cambridge University Press, pp.193–243.

Bower, G.H. (1967). A Multicomponent Theory of the Memory Trace. In K.W. Spence & J.T. Spence (Eds.), *The Psychology of Learning and Motivation*, Vol. 1, Academic Press, pp.230–325.

Brewer, W.F. and Treyens, J.C. (1981). Role of Schemata in memory for places. *Cognitive Phychology*, Vol. 13, pp.207–230.

Brewer, W.F. (1986). What is Autobiographical Memory? In D.C. Rubin (Ed.), *Autobiograpchical Memory*, New York: Cambridge University Press, pp.25–49.

Brown, N.R. and Schopflocher, D. (1998a). Event Clusters: An Organization of Personal Events in Autobiographical Memory. *Psychological Science*, Vol. 9, pp.470–475.

Brown, N.R. and Schopflocher, D. (1998b). Event Cueing, Event Clusters, and the Temporal Distribution of Autobiographical Memories. *Applied Cognitive Psychology*, Vol. 12, pp.305–319.

Bruner, J. (1987). Life as Narrative. *Social Research*, Vol. 54, pp.11–32.

Ceci, S.J. and Bronfenbrenner, U. (1985). "Don't Forget to Take the Cupcakes Out of the Oven": Prospective Memory, Strategic Time Monitoring, and Context. *Child Development*, Vol. 56, pp.152–164.

Ceci, S.J., Huffman, M.L.C., Smith, E. and Loftus, E.F. (1994a). Repeatedly Thinking aout a Non-event: Source Misattributions among Preschoolers. *Conciousness and Cognition*, Vol. 3, pp.388–407.

Ceci, S.J., Loftus, E.F., Leichtman, M.D. and Bruck, M. (1994b). The Possible Role of Source Missattributions in the Creation of False Beliefs among Preschoolers. *International Journals of Clinical and Experimental Hypnosis*, Vol. 42, pp.304–320.

Choudhury, T. and Pentland, A. (2003). Sensing and Modeling Human Networks using the Sociometer. *Proceedings of 7th IEEE International Symposium on Wearable Computers (ISWC2003)*, pp.216–222.

Clarkson, B. and Pentland, A. (1999). Unsupervised Clustering of Ambulatory Audio and Video. *Proceedings of the International Conference of Acoustics, Speech and Signal Processing.*

Clarkson, B. and Pentland, A. (2000). Framing through Peripheral Perception. *Proceedings of the 2000 IEEE International Conference on Image Processing (ICIP2000)*, pp.10–13.

Clarkson, B. and Pentland, A. (2001). Predicting Daily Behavior via Wearable Sensors. *MIT Technical Report Vismod TR451.*

Cohen, G. (1996). Memory in the Real World (2nd ed.) Hove, East Sussex: Psychology Press.

Conway, M.A. (1992). A Structural Model of Autobiographical memory. In M.A. Conway, D.C. Rubin, H Spinnler, & W.A. Wagenaar (Eds.), *Theoretical Perspectives on Autobiograpchical Memory*, Dordrecht: Kluwer Academic Publishers, pp.167–193.

Conway, M.A. and Rubin, D.C. (1993). The Structure of Autobiographical Memory. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory*, Hove, East Sussex: Lawrence Erlbaum Associates, pp.103–137.

Conway, M.A. (1996a). Autobiograpchical Knowledge and Autobiograchical Memories. In D.C. Rubin (Ed.), *Remembering Our Past: Studies in Autobiograpchical Memory*, New York: Cambridge University Press, pp.67–93.

Conway, M.A. (1996b). Autobiograpchical Memory. In E.L. Bjork & R.A. Bjork (Eds.), *Memory*, San Diego: Academic Press, pp.165–194.

Conway, M.A. and Pleydell-Pearce, C.W. (2000). The Construction of Autobiograpchical Memories in the Self-memory System. *Psychological Review*, Vol. 107, pp.261–288.

Csikszentmihalyi, M. and Beattie, O.V. (1979). Life Themes: A Theoretical and Empirical Exploration of Their Origins and Effects. *Journal of Humaistic Psychology*, Vol. 19, pp.45–63.

Czerwinski, M. and Horvits, E. (2002). An Investigation of Memory for Daily Computing Events. *Proceedings of the 16th Annual Human-Computer Interaction Conference*, pp.230–245.

DeVaul, R.W., Pentland, A. and Corey, V.R. (2003). The Memory Glasses: Subliminal vs. Overt Memory Support with Imperfect Information. *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC2003)*, pp.146–153.

Einstein, G.O., McDaniel, M.A., Smith, R.E. and Shaw, P. (1998). Habitual Prospective Memory and Aging: Remembering Intentions and Forgetting Actions. *Psychological Science*, Vol. 9, pp.284–288.

Eldelyi, M.H. (1996). *The Recovery of Unconscious Memories: Hypermnesia and Reminiscence*, Chicago: The University of Chicago Press.

Farringdon, J. and Oni, Y. (2000). Visual augmented memory (VAM). *Proceedings of 4th IEEE International Symposium on Wearable Computers (ISWC2000)*, pp.167–168.

Flexser, A. and Tulving, E. (1978). Retrieval Independence in Recognition and Recall. *Psychological Review*, Vol. 85, pp.153–171.

Geiselman, R.E., Fisher, R.P., Firstenberg, I., Hutton, L.A., Sullivan, S., Avetissian, I. and Prosk, A. (1984). Enhancement of Eyewitness Memory: An Empirical Evaluation of the Cognitive Interview. *Journal of Police Science and Administration*, Vol. 12, pp.74–80.

Gergen, K.J. and Gergen, M.M. (1988). Narrative and the Self as Relationship. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Vol. 21, San Diego: Academic Press, pp.17–56.

Gevers, T. and Smeulders, W.M. (1999). Colr Based Object Recognition. *Pattern Recognition*, Vol. 32, pp.453–464.

Gevers, T. (2001). Robust Histogram Construction from Color Invariants. *Proceedings of the 8th International Conference on Computer Vision*, pp.615–621.

Gelgon, M. and Thlhou, K. (2002). Automated Multimedia Diaries of Mobile Device Users Need Summarization. *Proceedings of the 4th International Symposium on Human Computer Interation with Mobile Devices and Services*, pp.36–44.

Goschke, T. and Kuhl, J. (1993). Representation of Intentions: Persisting Activation in Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 19, pp.1211-1226.

Goschke, T. and Kuhl, J. (1996). Remembering What to Do: Explicit and Implicit Memory for Intentions. In M.A. Brandimonte, G.O. Einstein, & M.A. McDaniel (Eds.), *Prospective Memory: Theory and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates. pp.53-91.

Hamasaki, M. and Takeda, H. (2003). Find better friends? - Reconfiguration of Personal Networks by the Neihborhood Matchimaker Method -. *Proceedings of the International Workshop on Semantic Web Foundations and Application Technologies (SWFAT2003)*, pp.73–76.

Healey, J. and Picard, R.W. (1998). StartleCam: A Cybernetic Wearable Camera. *Proceedings of the 2nd International Symposium on Wearable Computers*, pp42–49.

Heckhausen, H. and Beckman, J. (1990). Intentional Action and Action Slips. *Psychological Review*, Vol. 97, pp.36–48.

Hoisko, J. (2000). Context Triggered Visual Episodic Memory Prosthesis. *Proceedings of the 4th IEEE International Symposium on Wearable Computers (ISWC2000)*, pp.185–186.

Hyman, I.E.Jr., Husband, T.H. and Billings, F.J. (1995). False Memories of Childhood Experiences. *Applied Cognitive Psychology*, Vol. 9, pp.181–197.

Hyman, I.E.Jr. and Pentland, J. (1996). The Role of Mental Imagery in the Creation of False Childhood Memories. *Journal of Memory and Language*, Vol. 35, pp.101–117.

Hyman, I.E.Jr. and Billings, F.J. (1998). Individual Differences and the Creation of False Childhood Memories. *Memory*, Vol. 6, pp.1–20.

Intons-Peterson, M.J. (1996). Memory aids. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M.K. Johnson (Eds.), *Basic and Applied Memory*

Research, Vol. 2, *Practical Applications*, Hahwah, NJ: Lawrence Erlbaum Associates, pp.317–331.

Jang, S and Woo, W. (2003). Ubi-UCAM: A Unified Context-Aware Application Model. Proceedings of the 4th International and Interdisciplinary Conference on Modeling and Using Context, pp.178–189.

Jansari, A. and Parkin, A.J. (1996). Things that Go Bump in Your Life: Explaining the Reminiscece Bump in Autobiographical Memory. *Psychology and Aging*, Vol. 11, pp.85–91.

Jebara, T., Schiele, B., Oliver, N. and Pentland, A. (1998). DyPERS: Dynamic Personal Enhanced Reality System. *MIT Media Laboratory, Perceptual Computing Technical Report #463.*

Kato, T., Kurata, T. and Skaue, K. (2002a). VizWear-Active: Towards a Functionally-Distributed Architecture for Real-Time Visual Tracking and Context-Aware UI. *Proceedings of the 6th IEEE International Symposium on Wearable Computers (ISWC2002)*, pp.162–163.

Kato, T., Kurata, T. and Skaue, K. (2002b). Face Registration Using Wearable Active Vision System for Augmented Memory. *Proceedings of Digital Image Computing Techniques and Applications (DICTA2002)*, pp.252–257.

Kawashima, T., Nagasaki, T. and Toda, M. (2002). Information Summary Mechanism for Episode Recording to Support Human Activity. *Proceedings of the International Workshop on Pattern Recognition and Understanding for Visual Information Media*, pp.49–56.

Kawamura, T., Kono, Y. and Kidode, M. (2001a). Advanced Computing and Commucating Techniques for Wearable Information Playing -Basic Experiments for Video Diary Functions-. *Proceedings of the 29th Annual Conference of IIEEJ*, pp.85–92. (in Japanese)

Kawamura, T., Kono, Y. and Kidode, M. (2001b) A Novel Video Retireval Method to Support a User's Recollection of Past Events Aiming for Wearable Information Playing. *Proceedings of the 2nd IEEE Pacific-Rim Conference on Multimedia*, pp.24–32.

Kawamura, T., Ueoka, T., Ukita, N., Kono, Y. and Kidode, M. (2002a). Toward a Memory-aid System in the Wearable Information Playing Project. *Proceedings of the 3rd Meeting for Youth Community 2002 (MYCOM2002)*, pp.139–146. (in Japanese)

Kawamura, T., Kono, Y. and Kidode, M. (2002b). Wearable Interfaces for a Video Diary: towards Memory Retrieval, Exchange, and Transportation. *Proceedings of 6th International Symposium on Wearable Computes (ISWC2002)*, pp.31–38.

Kawamura, T., Fukuhara, T., Takeda, H., Kono, Y. and Kidode, M. (2002c). Ubiquitous Memories: Wearable Interface for Computational Augmentation of Human Memory based on Real World Object. *Technical Report of NAIST Information Science #NAIST-IS-TR2002012.*

Kawamura, T., Fukuhara, T., Takeda, H., Kono, Y. and Kidode, M. (2003a). Ubiquitous Memories: Wearable Interface for Computational Augmentation of Human Memory based on Real World Object. *Proceedings of 4th International Conference on Cognitive Science (ICCS2003)*, pp.273–278.

Kawamura, T., Kono, Y. and Kidode, M. (2003b). Nice2CU: Managing a Person's Augmented Memory. *Proceedings of 7th IEEE International Symposium on Wearable Computers (ISWC2003)*, pp.242–243.

Kidd, C.D., Orr, R., Abowd, G.D., Atkeson, C.G., Essa, I.A., MacIntyre, B., Mynatt, E.D., Starner, T.E. and Newstetter, W. (1999). The Aware Home: A Living Laboratory for Ubiquitous Computing Research. *Proceedings of the 2nd International Workshop on Cooperative Buildings*, **Position Paper**.

Kidode, M. (2002). Design and Implementation of Wearable Information Playing Station. *Proc. 1st CREST Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing*, pp.1–5.

Kono, Y., Kawamura, T., Ueoka, T., Murata, S. and Kidode, M. (2003). SARA: A Framework for Augmented Memory Albuming Systems. *Proceedings of the 2nd CREST Whorkshop on Advanced Computing and Communicating Techniques for Wearable Information Playing*, pp.20–34.

Lamming, M. and Flynn, M. (1994). Forget-me-not: Intimate Computing in Support of Human Memory. *Proc. FRIENDS21: International Symposium n Next Generation Human Interface*, pp.125–128.

Levy, R.L. and Loftus, G.R. (1984). Compliance and Memory. In J.E. Harris & P.E. Morris (Eds.), *Everyday Memory, Actions, and Absentmindedness.* New York: Academic Press, pp.93–112.

Maeda, H., Hirata, T. and Nishida, T. (1997). CoMeMo: Costructing and Sharing Everyday Memory. *Proceedings of the 9th International Conference on Tools with Artificial Intelligence*, pp.23–30.

Mann, S. (1997). Wearable Computing: A First Step Toward Personal Imaging. *Computer*, Vol. 30(2), pp.23–30.

Mäntylä, T. and Sgaramella, T. (1997). Interrupting Intentions: Zeigarnik-like Effect in Prospective Memory. *Psychological research*, Vol. 60, pp.192–199.

Marsh, R.L, Hicks, J.L. and Landau, J.D. (1998). An Investivation of Everyday Prospective Memory. *Memory & Cognition*, Vol. 26, pp.633–643.

Marsh, R.L, Hicks, J.L. and Bryan, E.S. (1999). The Activation of Unrelated and Canceled Intentions. *Memory & Cognition*, Vol. 27, pp.320–327.

McDaniel, M.A., Robinson-Riegler, B. and Einstein, G.O. (1998) Prospective Remembering: Perceptually Driven or Conceptually Driven Processes? *Memory & Cognition*, Vol. 26, pp.121–134.

Michael, B., Hans, W.G. and Albrecht, S. (2001). MediaCups: Experience with Design and Use of Computer Augmented Everyday Artefacts. *Computer Networks*, 35(4), pp.401–409.

Murakami, H. (2002). Information Acquitition and Reorganization from the WWW by using Memory-Organizer. *Bulletin of Osaka City University Media Center*, Vol. 6, No. 5-6, pp.341–357.

Nelson, K. (1993). The Psychological and Social Origins of Autobiographical Memory. *Psychological Science*, Vol. 4, pp.7–14.

Nickerson, R.S. and Adams, M.J. (1979). Long-term Memory for a Common Object. *Cognitive Psychology*, Vol. 11, pp.283–307.

Norman, D.A. (1981). Categorization of Action Slips. *Psychological Review*, Vol. 88, pp.1–15.

Picard, R.W. and Healey, J. (1997). Affective Wearables. *Proceedings of the 1st IEEE International Symposium on Wearable Computers*, pp.90–97.

Reason, J. (1979). Action not as Planned: The Price of Automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of Consciousness*, Vol. 1, London: Academic Press, pp.67–89.

Reason, J. (1984). Absentmindedness and Cognitive Control. In J.E. Harris & E. Morris (Eds.), *Everyday Memory, Actions, and Absentmindedness*, London: Academic Press, pp.113–132

Reason, J. (1990). Human Error. Cambridge: Cambridge University Press.

Rekimoto, J., Ayatsuka, Y. and Hayashi, K. (1998). Augment-able Reality: Situated Communication through Physical and Digital Spaces. *Proceedings of the 2nd IEEE International Symposium on Wearable Computers (ISWC'98)*, pp.68–75.

Rekimoto, J. and Ayatsuka, T. (2000). CyberCode: Designing Augmented Reality Environments with Visual Tags. *Proceedings of Designing Augmented Reality Environment.*

Rhodes, B. (1997). The Wearable Remembrance Agent: a System for Augmented Memory. *Proceedings of the 1st International Symposium on Wearable Computers (ISWC'97)*, pp.123–128.

Rhodes, B. (2003). Using Physical Context for Just-in-Time Information Retrieval. *IEEE Transactions on Computers*, Vol. 52, No. 8, pp.1011–1014.

Robinson, J.A. (1986). Autobiographical Memory: AHistorical Prologue. In D.C. Rubin (Ed.), *Autobiobraphical Memory*, New York: Bambridge University Press, pp.19–24.

Robinson, J.A. and Swanson, K.L. (1990). Autobiographical Memory: The Next Phase. *Applied Cognitive Psychology*, Vol. 4, pp.321–335.

Robinson, J.A. (1992). First Experience Memories: Contexts and Functions in Personal Histories. In M.A. Conway, D.C. Rubin, H. Spinnler, & W.A. Wagenaar (Eds.), *Theoretical Perspectives on Autobiographical Memory*, Dordrecht: Kluwer Academic Publishers, pp.223–239.

Ross, M., McFarland, C. and Fletcher, G.J.O. (1981). The Effect of Attitude on the Recall of Personal Histories. *Journal of Personality and Social Psychology*, Vol. 40, pp.627–634.

Ross, M., McFarland, C., Conway, M. and Zanna, M.P. (1983). Reciprocal Relation between Attitudes and Behavior Recall: Commitiiting People to Newly Formed Attitudes. *Journal of Personality and Social Psychology*, Vol. 45, pp.257–267.

Ross, M. and Conway, M. (1986). Remembering One's Own Past: The Consturction of Personal Histories. In R.M. Sorrentino & E.T. Higgins (Eds.), *Handbook of Motivation and Cognition*, New York: Guilford Press, pp.122-144.

Ross, M. (1989). Relation of Implicit Theories to the Construction of Personal Histories. *Psychological Review*, Vol. 96, pp.341–357.

Ross, M. and Buehler, R. (1994). Creative Remembering. In U. Neisser & R. Fivush (Eds.), *The Remembering Self: Construction and Accuracy in the Self-narrative*, New York: Cambridge University Press, pp.205–235.

Rubin, D.C., Wetzler, S.E. and Nebes, R.D. (1986). Autobiographical Memory across the Lifespan. In D.C. Rubin (Ed.), *Autobiographical Memory*, New York: Cambridge University Press, pp.202–221.

Rubin, D.C. and Schulkind, M.D. (1997). The Distribution of Autobiographical Memories across the Lifespan. *Memory & Cognition*, Vol. 25, pp.859–866.

Rubin, D.C. (2000). Autobiographical Memory and Aging. In D.C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer*, Hove, East Susex: Psychology Press, pp.131–149.

Sauer, J., wastell, D.G. and Hockey, G.R.J. (1999). Multiple-task Performance on a Computer-simulated Life Support System during a Space Mission Simulation. *Acta Astronautica*, Vol. 44, pp.43–52.

Sauer, J. (2000). Prospective Memory: A Secondary Task with Promise. *Applied Ergonomics*, Vol. 31, pp.131–137.

Schacter, D.L. (1995). Memory Distortion: History and Current Status. In D.L. Schacter (Ed.), *Memory Distortion: How Minds, Brains, and Societies Reconstruct the Past*. Cambridge, MA: Harvard University Press, pp.1–43.

Schaefer, E.G., Kozak, M.V. & Sagness, K. (1998). The Role of Enactment in Prospective Remembering. *Memory & Cognition*, pp.1-43.

Schiele, B., Oliver, N., Jebara, T. and Pentland, A. (1999). An Interactive Computer Vision System DyPERS: Dynamic Personal Enhanced Reality System. *Proceedings of International Conference on Vision Systems*, pp.51–65.

Singer, J.A. and Salovey, P. (1993). The remembered Self: Emotion and Memory in Personality. New York: The Free Press.

Sugiyama, T., Yoshimura, S., Suzuki, R. and Sumi, H. (2002). A 1/4-inch QVGA Color Imaging and 3-D Sensing CMOS Sensor with Analog Frame Memory. *Proceedings of IEEE International Solid-State Circuit Conference*, pp.434–435 and 479.

Sumi, Y., Sakamoto, R., Nakao, K. and Mase, K. (2002). ComicDiary: Representing Individual Experieces in a Comic Style. *Proceedings of the 4th Annual Conference on Ubiquitous Computing (UbiComp2002)*, pp.16–32.

Sural, S., Qian, G. and Pramaic, S. (2002). A Histogram with Perceptually Smooth Color Transition for Image Retrieval. *Proceedings of 4th International Conference on Computer Vision, Pattern Recognition and Image Processing*, pp.664-667.

Taylor, J.R. (1982). An Introduction to Error Analysis. *University Science Books*, 1982.

Takatori, K. (1980) The Role of Communication in the Memory Process - a Comparative Study of Individual Recall and Cooperative Recall-. *Educational Psychology*, Vol. XXVIII, No. 2, pp.108–113. (in Japanese)

Toda, M., Nagasaki, T., Iijima, T. and Kawashima, T. (2003). Structural Representation of Personal Events. *Proceedings of the ISPRS International Workshop on Visualization and Animation of Reality-based 3D Models.*

Tran, Q.T. and Mynatt, E.D. (2003). What was I Cooking? Towards Deja Vu Displays of Everyday Memory. Technical Report #GIT-GVU-03-33.

Tulving, E. and Thomson, D.M. (1973). Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review*, Vol. 80, pp.352–373.

Ueda, T., Amagasa, T., Yoshikawa, M. and Uemura, S. (2002). A System for Retrieval and Digest Creation of Video Data based on Geographic Objects. *Proceedings of the 13th International Conference on Database and Expert Systems Applications*, pp.768–778.

Ueoka, R., Hirota, K. and Hirose, M. (2001). Wearable Computer for Experience Recording. *Proceedings of the 11th International Conference on Artificial Reality and Teleexistence.*

Ueoka, T., Kawamura, T., Kono, Y. and Kidode, M. (2003). I'm Here!: a Wearable Object Remembrance Support System. *Proceedings of the 5th International Symposium on Human Computer Interaction with Mobile Devices and Services*, pp.422–427.

Ukita, N., Terabe, A., Kono, Y. and Kidode, M. (2002). Wearable Virtual Tablet: Fingertip Drawing Interface using an Active-Infrared Camera. *Proceedings of IAPR Workshop on machine Vision Applications*, pp.98–101.

Vortac, O.U., Edwards, M.B., Fuller, D.K. and Manning, C.A. (1993). Automation and Cognition in Air Traffic Control: An Empirical Intvestigation. *Applied Cognitive Psychology*, Vol. 7, pp.631–651.

Vortac, O.U., Edwards, M.B. and Manning, C.A. (1995). Functions of External Cues in Prospective Memory. *Memory*, Vol. 3, pp.201–219.

Vortac, O.U., Barile, A.L., Albright, C.A., Truitt, T.R., Manning, C.A. and Bain, D. (1996). Automation of Flight Data in Air Traffic Control. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M.K. Johnson (Eds.), *Basic and Applied Memory Research*, Vol. 2, *Practical Applications*, Mahwah, NJ: Lawrence Erlbaum Associates, pp.353–366.

Watts, D.J. and Strgatz, S.H. (1998). Collective Dynamics of 'Small-world' networks. *Nature*, 393, pp.440–442.

Weiser, M. (1991). The Computer for the 21th. *Scientific American*, 265(3), pp.94–104.

West, R. and Craik, F.I.M. (1999). Age-related Decline in Prospective Memory: The Roles of Cue Accessibility and Cue Sensitivity. *Psychology and Aging*, Vol. 14, pp.264–272.

Wilkins, A.J. and Baddeley, A.D. (1978). Remembering to Recall in Everyday Life: An approach to absentmindedness. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical Aspects of Memory*, New York: Academic Press, pp.27–34.

Wong, B., Starner, T. and McGuire, R.M. (2002). Towards Conversational Speech Recognition for a Wearable Computer Based Appointment Scheduling Agent. *Tech. Report*, 02-17, Georgia Tech., Dept, Grachics, Visualization, and Usability.

Publications

Journals

1. **Tatsuyuki Kawamura**, Tomohiro Fukuhara, Satoshi Murata, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kidode: Ubiquitous Memories: Associating Everyday Memory with Real World Objects using a Touching Operation, *Journal of the Institute of Electronics, Information and Communication Engineers*, D-I, 2005. (accepted) (in Japanese)

2. Takahiro Ueoka, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: I'm Here!: Wearable Interface System for Improving Efficiency of Finding Objects, *Journal of the Human Interface Society*, Vol.6, No.3, pp.19–29, 2004.

Book

1. **Tatsuyuki Kawamura**, Takahiro Ueoka, Yasuyuki Kono, Masatsugu Kidode: Wearable and Ubiquitous Video Data Management for Computational Augmentation of Human Memory, *Sagarmay Deb (ed) Video Data Management and Information Retrieval*, IRM Press, pp.33–76, ISBN1-59140-571-8, 2004.

International Conferences

1. **Tatsuyuki Kawamura**, Takahiro Ueoka, Yutaka Kiuchi, Yasuyuki Kono, Masatsugu Kidode: IPH: an Integrated Probabilistic Histogram for Measuring Image Similality in HSV Color Space, *Proceedings of the 9th IAPR Conference on Machine Vision Applications (MVA2005)*, Tsukuba, Japan, May 16–18, 2005. (accepted)

2. **Tatsuyuki Kawamura**, Toshiyuki Amagasa, Kenji Hatano, Kenji Hisazumi, Teruaki Kitasuka, Satoshi Murata, Takahiro Ueoka, Nirimichi Ukita, Yasuyuki Kono, Jun Miyazaki, Tsuneo Nakanishi, Akira Fukuda, Shunsuke, Uemura, Masatsugu Kidode: What is Augmented Memory? - Concept, Design, and Technology -, *Proceedings of the 3rd CREST/ISWC Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing*, pp.27–45, October 31, Arlington, VA, USA, 2004.

3. Takahiro Ueoka, **Tatsuyuki Kawamura**, Shigeyuki Baba, Shinichi Yoshimura, Yasuyuki Kono, Masatsugu Kidode: Wearable Camera Device for Supporing Object-Triggered Memory Augmentation, *Proceedings of the 3rd CREST/ISWC Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing*, pp.46–53, October 31, Arlington, VA, USA, 2004.

4. Satoshi Murata, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: Supporting On-Demand Experience Segmentation in the Ubiquitous Memories Environment, *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communications (RO-MAN2004)*, Okayama, Japan, September 20–22, 2004.

5. Takahiro Ueoka, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: Functional Evaluation of a Wearable Object Remembrance Support System, *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME2004)*, Taipei, Taiwan, June 27–30, 2004.

6. **Tatsuyuki Kawamura**, Takahiro Ueoka, Yasuyuki Kono, Masatsugu Kidode: Relational Analysis among Experiences and Real World Objects in the Ubiquitous Memories Environment, *Proceedings of the Pervasive 2004 Workshop on Memory and Sharing of Experiences*, pp.79–85, Vienna, Austria, April 20, 2004.

7. Yasuyuki Kono, **Tatsuyuki Kawamura**, Takahiro Ueoka, Satoshi Murata, Masatsugu Kidode: Real World Objects as Media for Augmenting Human Memory, *Proceedings of the Workshop on Multi-User and Ubiquitous User Interfaces (MU3I)*, pp.37–42, Funchal, Madeira Island, Portugal, January 13, 2004.

8. **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: Nice2CU: Managing a Person's Augmened Memory, *Proceedings of the 7th International Symposium on Wearable Computers (ISWC2003)*, pp.242–243, White Plains, NY, USA, October 21–23, 2003.

9. Takahiro Ueoka, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: I'm Here!: a Wearable Object Remembrance Support System, *Proceedings of the 5th International Symposium on Human Computer Interaction with Mobile Devices and Services (MobileHCI03)*, pp.422–427, Udine, Italy, September 8–11 2003.

10. **Tatsuyuki Kawamura**, Tomohiro Fukuhara, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kidode: Ubiquitous Memories: Wearable Interface for Computational Augmentation of Human Memory based on Real World Objects, *Proceedings of the 4th International Conference on Cognitive Science (ICCS2003)*, pp.273–278, Sydney, Australia, July 13–17, 2003.

11. Yasuyuki Kono, **Tatsuyuki Kawamura**, Takahiro Ueoka, Satoshi Murata, Masatsugu Kidode: SARA: A Freamwork for Augmented Albuming Systems, *Proceedings of 2nd CREST Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing*, pp.30–34, Nara, Japan, May 23–24, 2003.

12. **Tatsuyuki Kawamura**, Norimichi Ukita, Yasuyuki Kono, Masatsugu Kidode: HySIM: A Hybrid-space Image Matching Method for High Speed Location-based Video Retrieval on a Wearable Computer, *Proceedings of 2002 IAPR Workshop on Machine Vision Applications (MVA2002)*, pp.94-97, Nara, Japan, December 11–13, 2002.

13. Hideaki Takeda, Kazunori Terada, **Tatsuyuki Kawamura**: Artifact Intelligence: Yet Another Approach for Intelligent Robots, *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN2002)*, Berlin, Germany, September 25-27, 2002.

14. **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: Wearable Interfaces for a Video Diary: towards Memory Retrieval, Exchange, and Transportation, *Proceedings of the 6th International Symposium on Wearable Computers (ISWC2002)*, pp.31–38, Seattle, Washington, USA, October 7–10, 2002.

15. **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: A Novel Video Retrieval Method to Support a User's Recollection of Past Events for Wearable Information Playing, *Proceedings of the Second IEEE Pacific-Rim Conference on Multimedia (PCM2001)*, pp.24–31, Beijing, China, October 24-26, 2001.

Domestic Conferences (in Japanese)

1. Takahiro Ueoka, **Tatsuyuki Kawamura**, Shigeyuki Baba, Shinichi Yoshimura, Yasuyuki Kono, Masatsugu Kidode: A Wearable Color Camera to Support an Object Finding Task in Everyday Life, *Proceedings of the Interaction2005*, Tokyo, Japan, February 28 – March 1, 2005. (to appear)

2. Yutaka Kiuchi, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: Discriminating Operations to an Object with Unit Action Matching of Finger Motions, *Proceedings of the IPSJ-ISGCHI2004-HI-111, pp.115–122, Kagoshima, Japan, November 11–12, 2004.*

3. **Tatsuyuki Kawamura**, *Chin-Chan Lin, Yasuyuki Kono, Masatsugu Kidode: Nice2CU: Human-triggered Remembering Management System and Real World Interfaces,* Proceedings of the JSAI2004, *Ishikawa, Japan, May 31 – June 4, 2004.*

4. *Satoshi Murata,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Analysis of Time-series Data from Wearable Multi-sensors for Segmentation of Experience Video Data,* Proceedings of JSAI2004, *Ishikawa, Japan, May 31 – June 4, 2004.*

5. *Chin-Chan Lin,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Collection of Real World Information and Construction of Augmented Memories for a Communication-aid among Users,* Proceedings of the 66th Annual Conference of IPSJ 2004, *Vol.4, pp.229–230, Tokyo, Japan, June 27–30, 2004.*

6. **Tatsuyuki Kawamura**, *Takahiro Ueoka, Yutaka Kiuchi, Yasuyuki Kono, Masatsugu Kidode: Probabilistic Hue Histogram Integrated with Saturation and Value,* Proceedings of the 15th IPSJ-SIG-VIR, *pp.49–52, Osaka, Japan, October 31, 2003.*

7. *Satoshi Murata,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Toward on Demand Segmentation of Video Data of Experience using*

Time-Series Boilogical and Motion Information of Multi-Sensors, Proceedings of the 17th Annual Conference of JSAI 2003, *Nigata, Japan, June 23–27, 2003.*

8. **Tatsuyuki Kawamura**, *Tomohiro Fukuhara, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kidode: Interaction for Associating Real World Objects with Video Data of Experience,* Proceedings of 20th Conference of Japanene Cognitive Science Society (JCSS2003), *pp.240–241, Tokyo, Japan, June 6–8, 2003.*

9. **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Nice2CU: Managin Augmented Memory and Dynamic Net of Person's Data,* Proceedings of the 4th Meeting for Youth Community 2003 (MYCOM2003), *pp.23–30, Shiga, Japan, May 26–27, 2003.*

10. *Takahiro Ueoka,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: I'm Here! - A Wearable Object Remembering Support System,* Proceedings of the 65th Annual Conference of IPSJ 2003, *Vol.5, pp.179–182, Tokyo, Japan, March 25–27, 2003.*

11. *Satoshi Murata,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Implementation of Augmented Memory transportation and Sharing Functions in Ubiquitous Memories System,* Proceedings of the 65th Annual Conference of IPSJ 2003, *Vol.5, pp.187–190, Tokyo, Japan, March 25–27, 2003.*

12. *Satoshi Murata,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: Implementation of Augmented Memory Management,* Proceedings of Interaction 2003, *pp.57–58, Tokyo, Japan, February 27–28, 2003.*

13. *Yasuyuki Kono,* **Tatsuyuki Kawamura**, *Takahiro Ueoka, Satochi Murata, Norimichi Ukita, Masatsugu Kidode: Towards the Wearable Video Diary - Memory Retrieval, Editing, Transportation , and Exchange in Daily Life,* Proceedings of the IEICE PRMU2002-178, *pp.55-60, Nara, Japan, January 16–17, 2003.*

14. *Satoshi Murata,* **Tatsuyuki Kawamura***, Yasuyuki Kono, Masatsugu Kidode: Sharing Visual Augmented Memories in Ubiquitous Memories System,* Proceedings of the IPSJ-SIGCHI2002-HI-100*, pp.15–22, Hyogo, Japan, September 20–21, 2002.*

15. **Tatsuyuki Kawamura***, Takahiro Ueoka, Norimichi Ukita, Yasuyuki Kono, Masatsugu Kidode: Toward a Memory Aid System in the Wearable Information Playgin Project,* Proceedings of the 3rd Meeting for Youth Community 2002 (MYCOM2002)*, Shiga, Japan, June 17–18, 2002.*

16. **Tatsuyuki Kawamura***, Hideaki Takeda, Kazunori Terada, Tomohiro Fukuhara, Masaki Chikama, Kingo Koshiishi, Takahiro Ueoka, Masahiro Hamasaki: AgentBox: a Hardware for a Novel Interaction between Human and Artifacts,* Proceedings of the 16th Annual Conference of JSAI*, Tokyo, Japan, May 28–31, 2002.*

17. **Tatsuyuki Kawamura***, Tomohiro Fukuhara, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kidode: A Wearable System for Sharing Ubiquitous Memories in the Real World,* Proceedings of Interaction 2002*, pp.65–66, Tokyo, Japan, March 6–7, 2002.*

18. *Takahiro Ueoka,* **Tatsuyuki Kawamura***, Norimichi Ukita, Yasuyuki Kono, Masatsugu Kidode: An Object Remembrance Support System using a Wearable Camera System,* Proceedings of Interaction 2002*, pp.63–64, Tokyo, Japan, March 6–7, 2002.*

19. *Takahiro Ueoka,* **Tatsuyuki Kawamura***, Yasuyuki Kono, Masatsugu Kidode: Basic Study for a Wearable Object Registration and Retrieval System,* Proceedings of ITE Technical Report*, Vol.26, No.7, pp.25–30, Tokyo, Japan, January 25, 2002.*

20. **Tatsuyuki Kawamura***, Yasuyuki Kono, Masatsugu Kidode: Advanced Computing and Communicating Techniques for Wearable Information Playing - Basic Experiments for Video Diary Functions -,* Proceedings of the

29th Annual Conference of IIEEJ, *pp.85–92, Hokaido, Japan, June 21–22, 2001.*

21. **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: A Video Retrieval Method with Motion Information - Toward an Augmented Memory System for Wearable Information Playing -,* Proceedings of the 15th Annual Conference of JSAI, *Shimane, Japan, May 22–25, 2001.*

22. *Tomohiro Fukuhara,* **Tatsuyuki Kawamura**, *Fumihiro Matsumoto, Tooru Takahashi, Kazunori Terada, Takeshi Matsuzuka, Hideaki Takeda: Ubiquitous Memories: Human Memory Support System Using Physical Objects,* Proceedings of the 15th Annual Conference of JSAI, *Shimane, Japan, May 22–25, 2001. (Best presentation award)*

23. *Akihiro Terabe, Takahiro Ueoka,* **Tatsuyuki Kawamura**, *Yasuyuki Kono, Masatsugu Kidode: A New Input Interface for Wearable Computing,* Proceeding of the 15th Annual Conference of JSAI, *Shimane, Japan, May 22–25, 2001.*

Technical Reports

1. Takahiro Ueoka, **Tatsuyuki Kawamura**, Yasuyuki Kono, Masatsugu Kidode: View Angle Evaluation of a First-person Video to Support an Object-finding Task, *Information Science Technical Report #NAIST-IS-TR2005002*, Nara Institute of Science and Technology, 2005.

2. **Tatsuyuki Kawamura**, Chin-Chan Lin, Yasuyuki Kono, Masatsugu Kidode: Rating Human Relations for Recommendation of an Augmented Video Memory using Human Profile Data and Meeting Logs, *Information Science Technical Report #NAIST-IS-TR2003017*, Nara Institute of Science and Technology, 2003.

3. **Tatsuyuki Kawamura**, Takahiro Ueoka, Yutaka Kiuchi, Yasuyuki Kono, Masatsugu Kidode: Image Similarity Measurement by an Integrated Probabilistic Histogram, *Information Science Technical Report #NAIST-IS-TR2003015*, Nara Institute of Science and Technology, 2003.

4. Takahiro Ueoka, **Tatsuyuki Kawamura**, Norimichi Ukita, Yasuyuki Kono, Masatsugu Kidode: Remembering the Object Location with a Wearable Vision Interface, *Information Science Technical Report #NAIST-IS-TR2002017*, Nara Institute of Science and Technology, 2002.

5. **Tatsuyuki Kawamura**, Tomohiro Fukuhara, Hideaki Takeda, Yasuyuki Kono, Masatsugu Kidode: Ubiquitous Memories: Wearable Interface for Computational Augmentation of Human Memory based on Real World Objects, *Information Science Technical Report #NAIST-IS-TR2002012*, Nara Institute of Science and Technology, 2002.

Awards

1. IEEE Student Paper Award, the IEEE, February 2005.

   Tatsuyuki Kawamura, *Yasuyuki Kono, Masatsugu Kidode] Wearable Interfaces for a Video Diary: towards Memory Retrieval, Exchange, and Transportation, Proceedings of the 6th IEEE International Symposium on Wearable Computers (ISWC2002), pp.31–38, 2002.*

2. Best Presentation Award, the 15th Annual Conference of JSAI 2001, the Japanese Society for Artificial Intelligence (JSAI), June 2002.

   Tomohiro Fukuhara, **Tatsuyuki Kawamura**, *Fumihiro Matsumoto, Tooru Takahashi, Kazunori Terada, Takeshi Matsuzuka, Hideaki Takeda: Ubiquitous Memories: Human Memory Support System Using Physical Objects, Proceedings of the 15th Annual Conference of JSAI, 2001.*