Memory Ubiquitous: Providing Memories on Anything, Anywhere - A case study for cooking support -

Kazuaki Kondo, Masashi Kanegae, Takahiro Koizumi, Kanako Obata, and Yuichi Nakamura
Academic Center for Computing and Media Studies
Kyoto University, Kyoto, Japan
E-mail:kondo, kanegae, koizumi, obata, yuichi@media.kyoto-u.ac.jp

***Abstract*—This paper introduces a novel concept of "Memory Ubiquitous" which provides smart memory functions for appliances, facilities, and equipments used in everyday life. At the first step to realize it, we first propose three fundamental functions. These are capturing and recording the surrounding scene including human behaviors, annotating and editing the records to construct smart memories, and presenting the integrated memories. These functions can support human activities by aiding recall, giving instructions, and providing asynchronous communications. We chose a kitchen as an actual test bed and built a prototype cooking support system with the above smart memory functions. We implemented a context-sensitive record presentation based on state recognition and simultaneous display of multiple records that is adaptive to users' intentions. By evaluating the system's performance, we obtained several results that show how smart memory functions can assist human activities and how they can be extended for more useful services.**

***Keywords*-ambient memory, ubiquitous computing, user-centered design, cooking support**

I. Introduction

Our environment, which includes places, tools, and machines, has been storing our external memories for centuries. For example, a footprint indicates that someone has passed a place; traces of dirt from the hands on a tool indicate the area where someone often touches or holds it. The former also implies that the path leads to somewhere, and the latter implies that the tool is usually held in that manner. Such external memories play significant roles in our lives, even if we do not intend to leave behind such information. However, these environmental memories are not always reliable. The marks are not always self-descriptive; they are often ambiguous and they easily fade. Thus, we cannot always interpret them correctly.

Our research aims to realize more detailed and reliable environmental memories; we call this scheme "Memory Ubiquitous". With advanced media technologies, various artificial systems can be enhanced to capture people's behaviors and various situations while using them. These systems can then provide potential users with helpful information based on the recorded data. We can consider the following scenarios in a variety of locations.

- **HOME:** Smart memories in a kitchen or a living room capture the residents' activities. For example, a smart kitchen that supports cooking and health care is realized by capturing and retrieving cooking techniques, cooking materials, and the required amount of seasoning. A family recipe and techniques of implementing it can be passed to children or grandchildren. Health care information could also be provided at home through a smart living room or a smart bedroom.
- **OFFICE:** Memory Ubiquitous transforms appliances or workspaces into self-descriptive facilities that can explain who used them, how, when, as well as describe other events. Users can learn about the facilities, statistics, and various users' habits. Smart memories can be a media for sharing a community's knowledge.
- **PUBLIC SPACE:** Similar to the office scenario, people can share knowledge about facilities, spaces, and other activities. The records can be more public and include asynchronous messages. Note that security and privacy are important in this scenario.

Although the capture-and-provide idea sounds fine, it has not actually been realized, yet. One of the reasons is that data just recorded in a simple way are redundant, and reviewing them is tiring work. Another reason is that recorded data are not always self-descriptive. It is often the case that we do not notice a knack until it is explained by an experts.

This research tackles those issues based on memory ubiquitous scheme and addresses the fundamental functions for making it useful in a variety of places and situations. We investigated the design of such functions and installed them in a kitchen to construct a smart cooking support system. Then, we conducted basic experiments to delineate which functions are necessary and how helpful to users.

II. Intelligent Support with Memory Ubiquitous

A. Issues for Reliable Support

First, we discuss the characteristics of memory ubiquitous to clarify the requirements for providing reliable support. A memory ubiquitous system enables users to refer to the recorded past events and get useful information for determining their actions. The system achieves this framework by recording scenes and replaying them to the user, as shown in Fig. 1. For example in the case of cooking support, a user can refer to a past cooking session and recall what was successful, learn how another person cooks, and learn a variety of cooking techniques. This kind of support is closely related to concepts of ubiquitous computing and ambient intelligence. A variety of services use these concepts, e.g., a ubiquitous home[1], [2], [3] monitors its residents and offers them support in their daily activities. Memory Ubiquitous is a framework for how these ubiquitous or ambient systems acquire data and present them to support users. In other words, actual scenes captured around those systems can be considered as good data resources for ubiquitous or ambient memory. On the other hand, existing ubiquitous systems or user supporting systems are using data such as user guides given beforehand, or statistics gathered on-line. In several conventional cooking support systems[4], [5], [6] are good examples, in which well designed cooking manuals are given to the systems, and they are presented to the users when they are necessary. Some ubiquitous system gives users the information necessary for the possible next action, which is suggested by statistics of the users behaviors[7], [8].

Compared with these conventional approaches, memory ubiquitous has the following characteristics.

- Obviously, the quantity and the variety of data increases as a system is used. We can also expect that the quality of data can be improved if the users performed well by referring to previous records. Know-how about the environment can be accumulated by a cycle of improvements.
- Memories accumulated in a target system are usually consistent with situations that occur around it. For example, recorded data contain the same physical position and structures of appliances or instruments, and numerous kinds of information about a particular situation are applicable to supporting similar situations.
- Recorded data include not only the instructional information but also additional data such as information on habits, mistakes, failures, and current conditions. This surrounding information is often necessary for actual work and is also enjoyable in many cases.
- Scenes captured in a simple way are not essentially good resources for user support, because these recorded data are redundant in most cases and sometimes lack important information. For a similar reason, it is difficult to provide reliable support by simply replaying a large amount of recorded data.

The first three items above describe the advantageous characteristics and the last one indicates the problems that need to be solved. Tran et. al. addressed an important and interesting scheme on how short-term external memory aids people with memory handicap by providing feedback on cooking activities[9]. However, the problems on long-term memory as addressed in paper have not been considered.

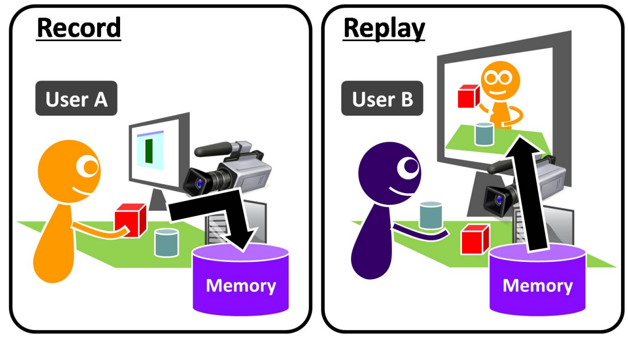

Figure 1. User support model based on memory ubiquitous scheme.

B. Functions for Smart Memory Ubiquitous

One of the most principal problems on memory ubiquitous is that elementary recording and replay functions are not sufficient, and that smart memory should be constructed and appropriately provided to the user. Kinds of requirement might be satisfied for the purpose. But in this paper, we here assume following two fundamental requirements to design elementary memory ubiquitous functions and evaluate them to discuss their effect and their advanced functions for practical use.

- **Requirement1:** Accumulated memories have various types of useful information sufficient for user support.
- **Requirement2:** Memories, including helpful information that the user wants, are presented when desired.

We propose three functions that address the two requirements.

- **Function(A)** recognizes the state of a scene to present situation-synchronized memories that provide context-adaptive support. This function assumes that it is informative to provide records of situations similar to the user's current situation.
- **Function(B)** simultaneously presents multiple records. It assumes that this is more informative than presenting a single record, because it is difficult to accurately estimate the user's intention and choose one record that truly supports it.
- **Function(C)** helps users annotate records to enhance them. Capturing a user's inner states (e.g. notions, feelings, impressions, and intentions) by simple passive observation is difficult, even if these often may provide important information for user support.

Functions (A) and (B) address requirement 2, and function (C) addresses requirement 1.

We investigated the automatic replay of situation-synchronized memories, including the negative effect caused by recognition failure and how the difficulties in the estimation of the user's request can be reduced by evaluating functions (A) and (B), respectively. We designed simple

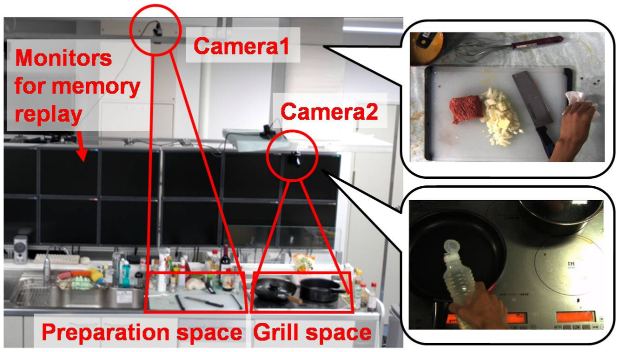

Figure 2. Overview of the prototype kitchen system and examples of captured memories.

algorithms assuming a particular support scenario to validate how the low-level implementations perform. We do not discuss function (C) due to space limitations. We determined some meaningful effects of an annotating assistance interface based on interactive approaches such as a virtual assistant framework [10].

III. Memory Ubiquitous in a Kitchen

A. System Overview

We assume a kitchen and cooking activity as a first test-bed of memory ubiquitous and support target, respectively. Since cooking activities are enough complicated situation and often appear in daily life, it must be a good case study for general applications. People often want kinds of help on their cooking activities, especially in the first time of cooking a new dish. Additionally, assistance for cooking activity includes common elements in other cases such as in office and public facilities. Therefore it is enough valuable to investigate how does the concept of memory ubiquitous support cooking activities. In this paper, the prototype system has particular memories about a target dish under an assumption that it knows a dish that a user plants to cook. This assumption is not so unreasonable because it can be satisfied by just asking him when he starts cooking.

The prototype system uses video clips as memories because they are easy to handle, contain considerable information, and can transmit nonverbal information. Figure 2 shows an overview of the prototype system, consisting of a kitchen, two cameras for observation, and monitors for display. The cameras capture a preparation space and a grill space, respectively, for both recording user behavior and recognizing cooking status. Additional equipments such as more cameras capturing a washing space, a refrigerating chamber, and behaviors of a user are available to acquire much information. Other media such as audio, force, and smell, could be also utilized.

B. Function(A): Providing Situation - Synchronized Memories

This function can be implemented by recognizing the scene and selecting corresponding memories. However, it still be a difficult problem to accurately recognize general situations including unexpected events from scenes and user's behaviors. In this paper, the proposed function recognizes scenes/situations into cooking step units with a combination of detected objects. Let an interval in which the same object combination emerges be a cooking step to recognize a situation in that unit. General cooking recipes and conventional cooking support systems also divide a sequence of a cooking operation into some cooking step for effective support. In this case, a recognition result is conducted from a data book that describes correspondences between a cooking step and a combination of emerged objects (Table I). For example, a "Cut onion" step corresponds to a combination of an onion and a knife in the preparation space. If the knife is used in only this step, a detection of the knife is enough for an unique recognition. We manually generated data books in current implementation. Automatic generation from a cooking recipe is probably possible. Detection and recognition of objects are based on their colors. Vivid color markers are attached to hard to recognize objects.

Implementation of the above function can be explained as following (see also Fig. 3). The recognition algorithm conducts a label L of a cooking step from objects O_1, O_2 that appear in each memory capturing cooking workspace (O_i includes "no object in the scene"). This processing is expressed as

$$L = f_L(O_1, O_2), \quad L = 1, 2, \cdots, N \tag{1}$$

where N is the number of cooking steps. f_L is determined by cooking procedures of a target dish as shown in Table I. At the recording phase, those recognition results divide a sequence of captured memory into smaller memories. Let an unit memory M_L be a piece of the divided memory corresponding to a cooking step L. It is expressed as

$$M_L = m_t, \quad t = (T_L, T_{L+1}) \tag{2}$$

where m and T_L mean a captured memory sequence and starting time of a cooking step L, respectively. After P times cooking has been done, accumulated memory M^P is expressed as

$$\{M^P\} = \{\{M_1^1, \cdots, M_N^1\}, \{M_1^2, \cdots\}, \cdots, \{M_1^P, \cdots\}\} \tag{3}$$

with an assembly of unit memories $M_i^j, i = 1, 2, \cdots N, j = 1, 2, \cdots P$.

At the supporting phase, a result of recognizing situation selects memories for display $\{M_{display}\}$ from $\{M^P\}$ as described by

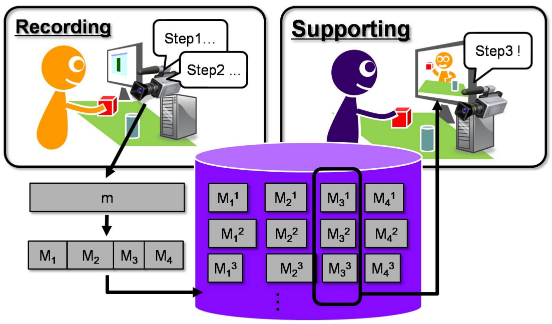

Figure 3. Labeling and selecting memories based on the recognizing situations.

Table I
CORRESPONDENCES BETWEEN COOKING STEPS AND OBJECTS IN THE SCENE FOR ONION-AND-MINCE OMELET.

Cooking step	Preparation space	Grill space
Cut onion	Knife, Onion	not considered
Stir egg	Bowl, Egg	not considered
SeasoningI	Salt, Pepper, Bowl	not considered
Prepare pan	not considered	Pan
Fry onion	not considered	Pan, Onion
Fry mince	not considered	Pan, Mince
SeasoningII	not considered	Pan, Salt, Pepper
Add egg and finish	not considered	Pan, Egg

$$\{M_{display}\} = \{M^1_{L_s}, M^2_{L_s}, \cdots, M^P_{L_s}\} \quad (4)$$

where L_s means a label of a recognized cooking step. $\{M_{display}\}$ consisting of multiple memories are simultaneously displayed with a spacial alignment addressed in the next section.

By the way, we usually consider what should we do at the next for sequential activities such as cooking. So it is natural to assume that information about near future situation could strongly assist the activities. Thus, we also implemented to present additional memories of the next cooking step as an expanded function. Presented records $\{M_{sync+}\}$ are expressed as

$$\{M_{sync+}\} = \{M^1_{S_c}, M^1_{S_c+1}, \cdots, M^P_{S_c}, M^P_{S_c+1}\}, \quad (5)$$

which differ from Eq. (2).

C. Function(B): Simultaneous Display of Multiple Memory Segments

A problem in simultaneous display of multiple videos is trade-off between physical size/resolution of a display device and resolutions/the number of videos for display. To say concretely, multiple movies displayed on a normal single monitor are hard to watch for users. Here we attack to this problem with using a tiled-display technique that allows to use multiple monitors as a single large display space. Flexible construction by monitor alignment and lower cost than other equipments that have similar capabilities are also merits of a tiled-display.

(a) (b)

Figure 4. Snapshots of the experiments.(a) Simultaneous display of multiple records. (b) With additional records of the next cooking step.

We constructed a tiled display consisting of 20-inch, 1920×1080-pixel, flat-panel monitors in a 6×2 (horizontal $\times$ vertical) array, and the control software SAGE. The selected records $\{M_{sync}\}$ are replayed on the tiled display; each record corresponds to one of the two horizontally contiguous monitors, as shown in Fig. 4. This method enables the users to understand the presence of multiple records, to browse information without active control, and to ignore the display when they do not need any information. Although we should also discuss issues such as how to align multiple records and how to reduce records for present that are too large for the display space, a simple presentation of a few records along with a recorded sequence is implemented here.

IV. EXPERIMENTAL VALIDATIONS

A. Cooking Support with the Prototype System

We conducted cooking support experiments to validate the performance of the implemented fundamental functions. Actually, we compared following three configurations.

- **Control:** Just repeat one continuous memory.
- **Proposed1:** Simultaneously display multiple situation-synchronized memories.
- **Proposed2:** Simultaneously display multiple situation-synchronized memories and near future memories.

The common experimental configurations are described below.

- The support target dish is an onion-and-mince omelet, which is simple but requires a little know-how to yield a tasty and attractive result.
- Three people ($p = 3$) cooked the omelet in advance with a recipe consisting of eight cooking steps ($n = 8$), as shown in Table I. Their behaviors were recorded to

Table II
THE QUESTIONNAIRES AND THEIR AVERAGED ANSWERS. "5" IS THE HIGHEST SCORE FOR 5-POINT SCALE QUESTIONS. "*" INDICATES THAT A QUESTION HAS NO MEANING BECAUSE NEAR FUTURE MEMORY DO NOT APPEAR IN PROPOSED 1 CONFIGURATION.

Term	Question	Type	Control	Proposed 1	Proposed 2
1	Did you get desired information from given support ?	5-point scale	2.0	4.0	4.2
2	Is desired information given when you want it ?	5-point scale	2.4	3.2	4.6
3	What is problem of the system ? Do you want any other functions ?	free form	-	-	-
4	Is there any informative content in the next step memories ?	5-point scale	3.0	*	4.8
5	Write what is informative content if you find ?	free form	-	-	-

construct memories based on the recognition results of function (A).

- The recorded data were enhanced by comments about the amount of ingredients, cooking procedures, and important points. We manually annotated the data without any intelligent support, which will be supported by the function (C).

In the experiments, five subjects who have variety in cooking frequency in daily life, are supported by the memory ubiquitous functions. Figure 4 shows snapshots of the experiments. The results were evaluated by two methods: the questionnaire survey listed in Table II for each subject after the experiment, and the observation of the behaviors and statements of the subjects, who were asked to describe their feelings, impressions, and thoughts during the experiments.

On the proposed 1 configuration, we obtained some positive results regarding function (A) as comments "I could consult the desired information with appropriate timing," and "Informative memories appeared when I hesitated," and also as some observed behaviors synchronized with the displayed data with statements such as "I will just do it." However, the scores shown in Table II indicate relatively poor performance. The main reason is that the system sometimes failed to recognize the cooking steps and presented unhelpful information that mismatched the current cooking progress. Function (A) is not very useful at the current implementation level. We also observed many interesting behaviors, such as emulating recorded operations similar to the subject's operation and consulting informative or interesting operations from the multiple presented memories produced by function (B). Even the current simple implementation successfully provides the requested information.

Furthermore the proposed 2 configuration improves the scores for both term 1 and term 2, as shown in Table II. The possible reasons are as follows.

- Because twice as much information is presented, the probability of the availability of the desired information increased.
- The presentation of multiple records that cover two cooking steps worked as a fail-safe for function (A) ; Even if the situation recognizing process fails to estimate a cooking step, the next step of the wrong estimation may match to it.

Displaying memories about the next cooking step itself also yielded good scores and positive comments such as "I understood how to proceed with the information given in advance," and "Memories about the next step seem to be informative and allow sufficient time for the cooking operations." We confirmed that the presentation of additional records about near-future situation greatly enhanced the performance. It is especially interesting that the small change in function (B) not only achieved the intended effect, but also improved the performance of function (A). We believe that the discovery of this synergy effect is important for constructing a smart memory ubiquitous system.

V. DISCUSSIONS

We acquired some positive perceptions and of course remaining issues through the experiments. Here we discuss them from the viewpoints of how the requirements listed in section II-B were satisfied by the proposed functions. Because detailed proposal of the function (C) was skipped in this paper, we do not discuss the requirement1 related it here.

For the requirement2, we proposed a straight-forward method ; display memories whose contents similar to the current situation with object recognition. In many cases, providing those memories in cooking step unit was informative. Equivalent performance can be expected in other cases if these operations can be divided into smaller one. Inappropriate support cases are categorized into two groups. One is that users need to wait desired information appearing for a long time. The other is failure of the situation recognizing, which causes recording memories with wrong labels and selecting irrelative memories. The former problem may be solved by dividing a target operation into finer steps. But an optimal method for the dividing is not obvious. Behaviors of users can be available to define and recognize each operation step, other than combinations of ingredients and cookwares. Amount of short memories that a person can remember discussed in section **??** seems to also be considerable. However fine division and complicated definition of each step tends to make recognition accuracy problems hard. Kinds of more efficient object recognition can attack to this trade-off than the current implementation based on color features of items in a particular recipe. We are now approaching an accurate object recognition method with user's collaboration for general scenes including unexpected events.

On the other hand, simultaneous display of multiple memories performs as a fail-safe for difficulty on accurate estimation of required information and select an unique memory for display. Furthermore it provides additional effects - multiple memories allow users to browse and compare them to select the best one, and give useful/interesting information that they have not thought so. Especially providing both of memories whose contents are synchronized to the current situation and a near future one conducts great performance. On the other hand, cognitive burden and physical constraints limit the number of memories for display. One approach is to reduce accumulated memories down to the number of being able to display and being easy to consult for users. This filtering should be determined by how fine information do users request and how accurately the requests are estimated. For example, when a user requests rough information or the estimation accuracy is low, it seems to be efficient to display memories including various contents to cover wide range of information. When he requests detailed information such as subtle movements or the estimation is reliable, slightly different memories focusing on a particular scene may be suitable. A temporal step-by-step concretion of memories like coarse-to-fine method may be a good approach when users can control a display interface. How to align multiple memories addressed in section III-C remains as an important issue. Additional display of guidelines for consulting memories is also worth to be implemented.

VI. Conclusion

We proposed a new concept, named "Memory Ubiquitous" for tools, appliances, and equipments used in everyday life. Such environment automatically record information, including the behavior of its users, and actively replay them. Two functions, providing situation-synchronized memories and simultaneous display of multiple records, were installed in a kitchen space and validated through cooking support experiments. The experimental results suggest that the proposed functions, especially additional records about near-future situations, are possible effective approaches to smart user support.We are now attempting to investigate an interactive interface that helps users annotate recorded memories. Recursive improvement of accumulated memories is an important function for constructing smart memory.

References

[1] J. M. Hsu, W. J. Wu and I. R. Chang, "Ubiquitous Multimedia Information Delivering Service for Smart Home", In Proc. of Int. Conf. on Multimedia and Ubiquitous Engineering, pp. 341–346, 2007.

[2] T. Yamazaki, "Ubiquitous Home: Real-Life Testbed for Home Context-Aware Service", In Proc. of Int. Conf. on Testbeds and Research Infrastructures for the Development of Networks and Communities, pp. 54–59, 2005.

[3] de Silva, C. Gamhewage, B. Oh, T. Yamasaki and K. Aizawa, "Experience retrieval in a ubiquitous home", In Proc. of the 2nd ACM workshop on Continuous archival and retrieval of personal experiences, pp. 35-44, 2005.

[4] A. Hashimoto, N. Mori, T. Funatomi, Y. Yamakata, K. Kakusho and M. Minoh, "Smart Kitchen: A User Centric Cooking Support System", In Proc. of Information Processing and Management of Uncertainty in Knowledge-Based Systems, pp. 848–854, 2008.

[5] K. Miyawaki, M. Sano, S. Yonemura and M. Matsuoka, "A cooking support system for people with higher brain dysfunction", In Proc. of the ACM multimedia 2009 workshop on Multimedia for cooking and eating activities, pp. 47–52, 2009.

[6] S. Murakami, T. Suzuki, A. Tokumasu and Y. Nakauchi, "Cooking Procedure Recognition and Support by Ubiquitous Sensors", Journal of Robotics and Mechatronics, Vol. 21, No. 4, pp. 498–506, 2009.

[7] H. Aoyama, M. Ozeki and Y. Nakamura, "Smart Cooking Support System based on Interaction Reproducing Model", In Proc. of ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities, pp. 39–45, 2009.

[8] G. D. Abowd and E. D. Mynatt, "Charting Past, Present and Future Research in Ubiquitous Computing", ACM Transactions on Computer-Human Interaction, Vol. 7, pp. 29–58, 2000.

[9] Q. T. Tran, G. Calcaterra and E. D. Mynatt, "Cook's Collage: Deja vu display for a home kitchen", In Proc. of Home-Oriented Informatics and Telematics, pp. 15–32, 2005.

[10] M. Ozeki, S. Maeda, K. Obata and Y. Nakamura, "Virtual Assistant: Enhancing Content Acquisition by Eliciting Information from Humans", Multimedia Tools and Applications, Vol. 44, No. 3, pp. 433–448, 2009.