Research and Practice in Technology Enhanced Learning
Vol. 3, No. 3 (2008) 297–312

A FRAMEWORK FOR CAPTURING, SHARING AND COMPARING LEARNING EXPERIENCES IN A UBIQUITOUS LEARNING ENVIRONMENT

HIROAKI OGATA*, TORU MISUMI, TSUYOSHI MATSUKA,
MOUSHIR M. EL-BISHOUTY and YNOEO YANO

Department of Information Science and Intelligent Systems
University of Tokushima, 2-1 Minamijosanjima
Tokushima, 770-8506, Japan
**ogata@is.tokushima-u.ac.jp*
http://www-b4.is.tokushima-u.ac.jp/ogata

This paper proposes a personal learning assistant called LORAMS (Link of RFID and Movies System), which supports learners with a system to share and reuse learning experiences by linking movies to environmental objects. We assume that every object has RFID tags and mobile devices have a RFID reader and can record a video anytime and anyplace. By scanng RFID tags of real objects, LORAMS can provide only video segments that include the objects. Also LORAMS recommends similar videos to be compared. In LORAMS, the video recording and RFID tagging are used purposely to support further teaching or learning rather than "just record it and use it in some day". We think that LORAMS can be applied to various kinds of domains that employ several kinds of real objects and vary the results depending on the combination of the objects; for example, cooking, checking upon cars such as oils, battery, and tires, surgery operations and chemical bioreactor experimentations. An evaluation was conducted in the context of cooking. The subjects could easily find the difference between their videos and expert's videos and improve how to cook fried rice.

Keywords: Computer supported ubiquitous learning; mobile learning; pervasive learning; RFID tag; video; SMIL.

1. Introduction

Ubiquitous computing (Abowd & Mynatt, 2000) will help organize and mediate social interactions wherever and whenever these situations might occur (Lyytinen & Yoo, 2002). Its evolution has recently been accelerated by improved wireless telecommunications capabilities, open networks, continued increases in computing power, improved battery technology, and the emergence of flexible software architectures (Sakamura & Koshizuka, 2005). With those technologies, CSUL (Computer Supported Ubiquitous Learning) is realized, where an individual and collaborative learning in our daily life can be seamlessly included.

One of the most important ubiquitous computing technologies is the RFID (radio frequency identification) tag, is a rewritable IC memory with non-contact communication facility. This cheap, tiny RFID tag makes it possible to tag almost everything, replaces the barcode, helps computers to be aware of their surrounding objects by themselves, and thereby detects some aspects of the user's context (Borriello, 2005). The features of RFID tag are as follows:

(1) Non line-of-sight reading: unlike the bar code, RFID does not require line-of-sight reading. In addition, the distance range for RFID reader is longer than bar code scanning range.
(2) Multiple tag reading: Unlike a bar code reader, RFID unit can read multiple tags at the same time. This feature enables counting a number of objects in a second. This is one reason that one of the key applications of RFID is supply-chain management.
(3) Data rewritable: RFID has a memory chip that can be rewritten using an RFID unit, on the other hand, the data of bar code is not changeable.
(4) High durability: Tags are very sturdy from vibrations, contamination (dust and dirt), and abrasion (wear). Hence, tags can be permanently used.
(5) Ease of maintenance: There are two types of RFID tags. One is passive, which does not use any battery. The power comes from the reader unit. Therefore, passive tags can be used permanently. The other one is active, which contains batteries and has a longer range than passive ones.

We assume that almost all products will be attached with RFID tags in the near future, where we will be able to learn at anytime at any place from every object by scanning its RFID tag. Also we think that almost mobile devices will have a video camera with record capabilities.

The fundamental issues of CSUL are

(1) How to capture and share learning experiences that happen at anytime and anyplace.
(2) How to retrieve and reuse them for learning.

As for the first issue, video recording with handheld devices will allow us to capture learning experiences. Also consumer generated media (CGM) services such as YouTube [http://www.youtube.com/] helps to share those videos. The second issue will be solved,by identifying objects in a video with RFID so that the system can recommend the videos in similar situations to the situation where the learner has a problem.

This paper proposes LORAMS (Linking of RFID and Movie System) for CSUL. There are two kinds of users in this system. One is a provider who records his/her experience into videos. The other is a user who has some problems and retrieves the videos. In this system, a user uses his/her own PDA with RFID tag reader and digital camera, and links real objects and the corresponding objects in a movie and shares it among other learners. Scanning RFID tags around the learner enables us to bridge the real objects and their information into the virtual world. LORAMS detects the objects around the user using RFID tags, and provides the user with the right information in that context.

In the previous work, we developed LORAMS using RealPlayer and applied the systems for computer assembling (Ogata, Matsuka, El-Bishouty & Yano, 2007). From the evaluation, RFID is very useful to identify objects precisely. However, we found that LORAMS system should support annotations for sharing knowledge and comparing videos. In this paper, LORAMS system has been improved in order to utilize the full advantage of RFID to share and compare personal experiences for ubiquitous learning.

2. Literature Review

This section describes related work on educational systems using RFID tags, life-log systems and video annotation systems.

2.1. *Educational systems using RFID tags*

There are two kinds of educational applications using RFID tags. The first type is applications that can identify the objects on a table and support face-to-face collaboration. For example, EDC (Envisionment and Discovery Collaboratory) (Arias *et al.*, 1999) and Caretta (Sugimoto, Hosoi & Hashizume, 2004) consist of a sensing board and objects with RFID tags such as house, school, etc. Detecting objects on the table enables the systems to show the simulation such as urban planning. Also TANGO (Tag Added learNinG Objects) system supports learning vocabularies (Ogata & Yano, 2004). The idea of this system is to stick RFID tags on real objects instead of sticky labels, annotate them (e.g. questions and answers), and share them among others. The tags bridge authentic objects and their information into the virtual world. The second type is applications that can detect the learner's location using RFID tags that allows the system to track the learner's positions and to send the right messages to the learner. eXspot (Hsi & Fait, 2005) is an example of this type of application. Designed for museum educators, it can capture the user's experiences at a museum for later reflection. This system consists of a small RFID reader for mounting on museum exhibits, and RFID tag for each visitor. While using RFID, a visitor can bookmark the exhibit s/he is visiting, and then the system records the visitor's conceptual pathway. After visiting the museum, the visitor can review additional science articles, explore online exhibits, and download hands on kits at home via a personalized web page. In contrast to those systems, LORAMS employs RFID tags to indentify objects in each scene of a video. Users can also find videos that include the objects by scanning tags.

2.2. *Life-log*

The idea of a "life-log" or personal digital archives is a notion that can be traced back at least 60 years (Bush, 1945). The idea is to capture everything that ever happened to us, to record every event we have experiences and to save every bit of information we have ever touched. For example, SenseCam (Hodges, Williams, Berry, Izadi, Srinivasan *et al.*, 2006) is a sensor augmented wearable still camera. Hodges *et al.* (2006) proposed to capture a log of the wearer's day by recording a series of images and capturing a log of sensor data. Reviewing this information will help the wearer recollect aspects of earlier experiences that have subsequently been forgotten, providing a powerful retrospective memory aid. MyLifeBits (Gemmell, Bell & Lueder, 2006) tores scanned material, e.g. articles, books, as well as digital data, e.g. emails, web pages, phone calls, and digital photos taken by SenseCam. While SenseCam employs a stills camera, LORAMS uses a video for capturing the experiences. Therefore, LORAMS helps to understand how to accomplish a task from a video, compared to SenseCam. Also, we can use a wearable (head-mounted) video camera with microphone instead of the digital video camera that we used in the experiment reported later in this paper.

The Ubiquitous Memory system is also a life-log system using a video and RFID tags (Kawamura, Fukuhara, Takeda, Kono, Koide, 2006). While this system provides video in response to touching only one object in order to recall the past memory about that object, LORAMS provides an advanced search function using multiple objects and allows users to make annotations and comparisons in order to help learning.

2.3. *Video annotations*

There are many related works that support manual text-annotation and keywords annotation of a video. (Davis, 1993; Nagao, Shirai & Squire, 2001; Smith & Lugeon, 2000). For example, the DIVER system is a web-based application that allows researchers to use their web browser to view digital video clips and add text annotations to specific points in space and time within the video (Lindgren, Pea, Lewis & Rosen, 2007). These methods need a lot of time to make annotations. Therefore, Yamamoto & Nagao (2005) proposed that the viewer puts the annotation into the video contents in order to decrease the load on the video provider. Also, artificial intelligent methodology was applied to give annotations automatically. In contrast to those systems, LORAMS does not require users to make manual text-annotations. By scanning RFID tags, the users can find suitable videos. Therefore, most of the subjects in our experiment stated that it was very easy to make RFID-tag based annotations into video.

3. Design and Implementation of LORAMS

3.1. *Features*

The characteristics of LORAMS are as follows:

(1) The learner's experience is recorded into a video and linked to RFID tags of real objects. The video can be shared with other learners. Therefore, it is not

necessary to add keywords or annotations into a video and is easy to make an index of the video to be shared with other learners.

(2) Learners can find suitable videos by scanning RFID tags of real objects around them without entering any keyword.
(3) Based on ratings by learners and the system, the results are listed.

There are three phases for LORAMS as follows:

(4) Video recording phase.
(5) Video search phase.
(6) Video replay and compare phase.

The video recording process requires a PDA, an RFID tag reader, a video camera, and wireless access to the Internet. First, a user has to start recording video at the beginning of the task. Before using objects, the user scans RFID tags and the system automatically sends the data and its time stamp to the server. After completing the task, the user uploads the video file to the server and the server automatically generates an SMIL (Synchronized Multimedia Integration Language) file to link the video and the RFID tags.

On the other hand, video search and reply processes need a PDA, an RFID tag reader, and a video player. The user scans RFID tags around him/her and/or enters keywords of the objects, and then the system sends them to the server and shows the list of the videos that match the objects and keywords. The system, moreover, extracts a part of the video that matches with these objects.

3.2. *User interface*

In the recoding phase, the user sets up the information on the RFID reader such as port number and code type, and enters the user name. When the user uses an object, s/he pushes the "start" button and scans the RFID of the object. Also, when the user finishes the work using the object, s/he pushes the "end" button and scans the RFID of the object. The RFIDs and the time stamps of the scans are sent to the server by pushing the "send" button. As shown in the right of Figure 1, the RFIDs are linked to the video. Users can create their own user id and password before using LORAMS, and a video file can be uploaded through the web page.

By entering keywords and/or scanning RFID tags of physical objects in (A) in Figure 2, video search starts. LORAMS starts to search videos and shows the video in the order of appropriateness. The list in (C) shows the videos that have been registered recently. By selecting a video from the list (B), the video playback window appears. The video title, the author's name, and the recorded date are shown in (D), and all the objects are listed in (E) in time order. By clicking an item in the list, the system jumps to the video segment that includes the selected item. Pictures of the items are shown in (F). By clicking one of the pictures, the system shows the video segments that include the selected item. By pushing a button in (G), the user can

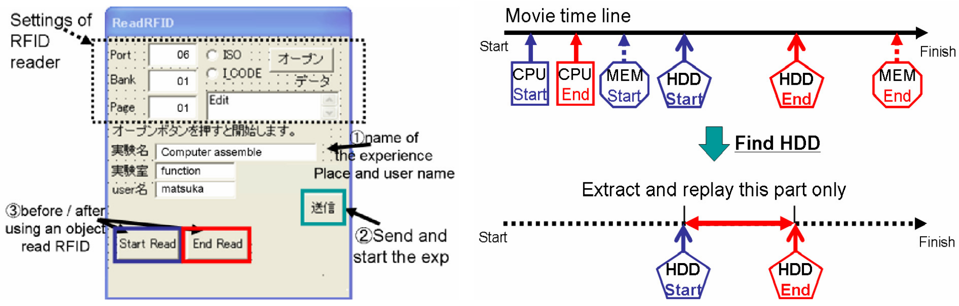

Figure 1. The interface of the recording phase (left) and video time line (right).

rate the video by the scale of 1 to 5. Playback (such as fast-forward) is controlled by the tool bar in (H). Videos similar to the playing video are listed in (I).

Figure 3 shows the interface for video annotation. The system has an annotation function for adding information on videos. Because not only a video provider but also a video viewer can make an annotation, a lot of information can be shared from different perspectives. We believe this information is useful for learning.

The system provides the following annotation function using the icons in the right side of the video window:

(A) User can insert a text into a video picture after stopping replay.
(B) User can add a title to a scene by selecting the time period.
(C) User can trim a scene by selecting the time period.
(D), (E), (F) User can insert an arrow into a video picture.
(G) As a memo, user can insert a URL of a web page, an image, a file into a video picture.

User can compare two different videos in the window as shown in Figure 4. For example, the left is a video of an expert, and the right is a video of the user after the user watched the expert's video. The tile of the video is shown in (A) and the video is replayed in (B). The timeline of the left video is shown in (D) and that of the right video is (E). In Figure 4, the user can find that the timeline (E) took longer time than (D) and the work efficiency of the user is not good. On the timeline, a colored rectangle shows an object that the user used at the timing. If the mouse pointer is over the colored rectangle, the system shows the picture of the corresponding object in (F). Since the same object has the same color, the user can easily recognize when the object was used in the two videos.

3.3. *System configuration*

We have developed LORAMS, which works on a Fujitsu Pocket Loox v70 with Windows Mobile 2003 2nd Edition, RFID tag reader/writer (OMRON V720S-HMF01),

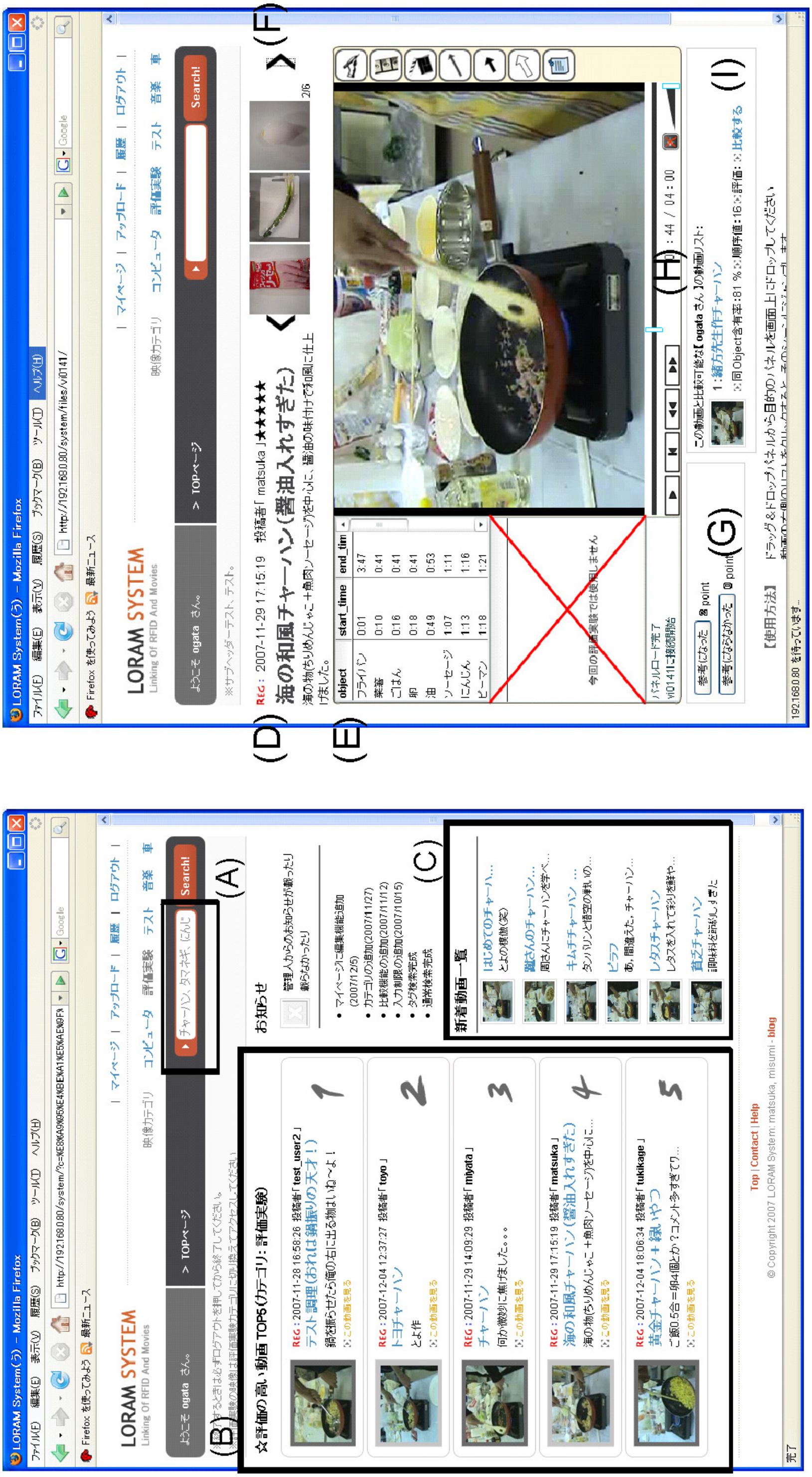

Figure 2. The interface of the video search (left) and video playback (right).

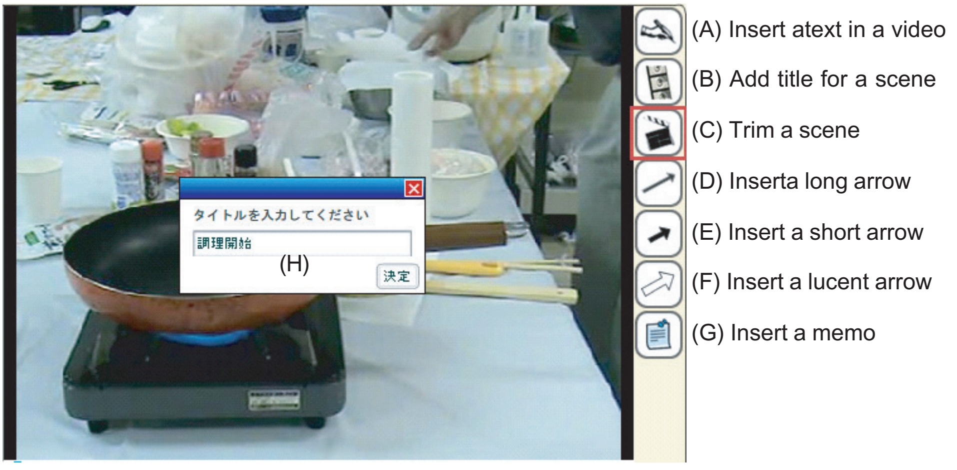

Figure 3. Annotation interface.

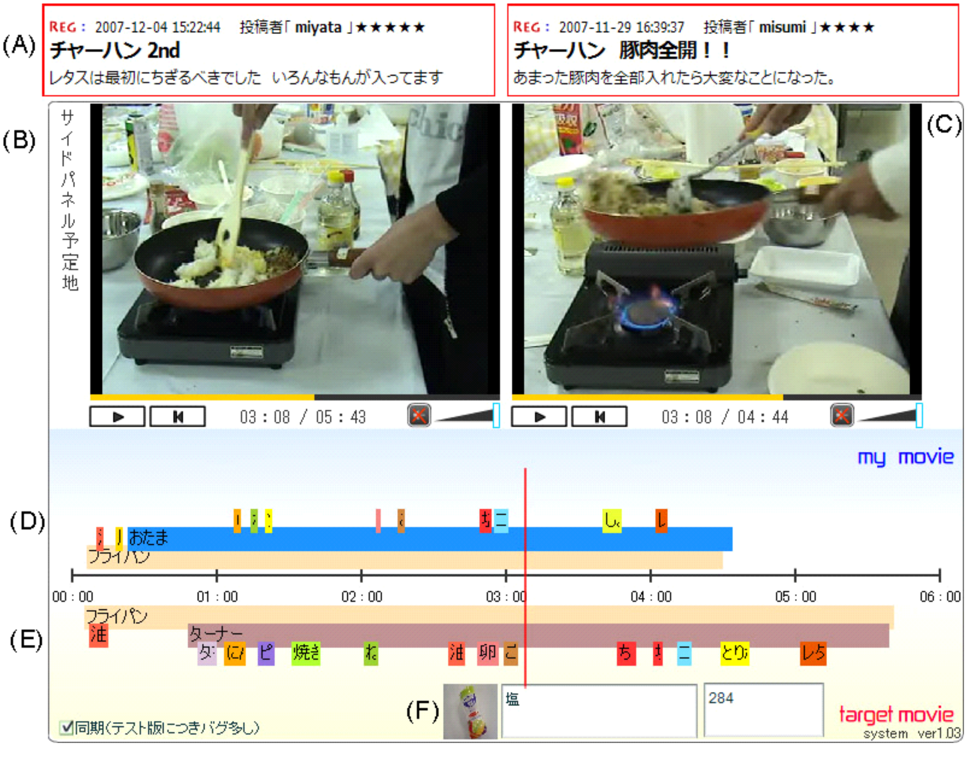

Figure 4. Interface for comparing two videos.

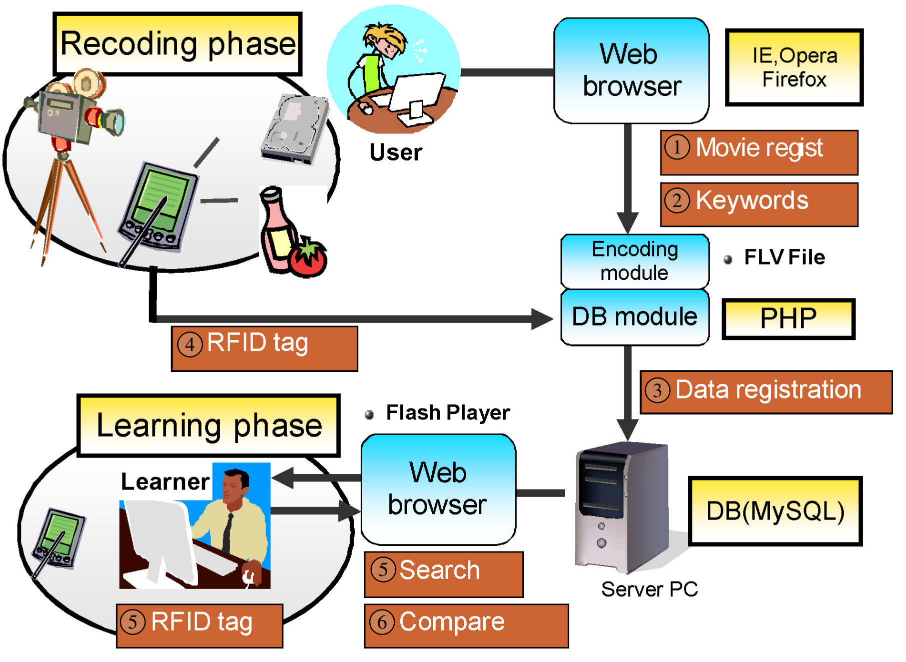

Figure 5. System configuration.

and WiFi (IEEE 802.11b) access. The reader/writer is attached on a CF (Compact Flash) card slot of PDA. The tag unit can read and write data into and from RFID tags within 5 cm distance, and it works with a wireless LAN at the same time. We used the short range of RFID reader because users should scan only the objects including the video clip. In other words, we did not use a long range RFID reader because it scans the objects that are not included in the video, for example, objects in the user's back.

The LORAMS program has been implemented using Embedded Visual C++ 4.0 and PHP 5.0. The video is played using Flash player. The server application consists of the following modules:

(1) Database entry: It stores the RFID reading time stamp into the DB.
(2) Database: This system uses My SQL server as a database.
(3) Database search: This module matches videos with keywords and RFID tags.
(4) SMIL generation: After finding the segments that contain the keywords and RFID tags, this module generates SMIL files for each segment.

3.4. *Ranking method*

There are two recommendation in LORAMS. First, the system recommends a user the right videos according to the objects scanned by the user. Second, the system recommends a user the right videos that the user should compare with the target video selected by the user.

(1) Video recommendation

A ubiquitous computing environment enables people to learn at any time and any place. The challenge in an information-rich world is not only to make information available to people at any time, at any place, and in any form, but specifically to say the right thing at the right time in the right way (Fischer, 2001). This system employs the following equation to rank the search results in order to recommend the right video:

$$l = \sum_{i=1}^{5} W_i X_i$$

where,

X_1: subjective value given by the provider and $0 \leq X_1 \leq 1$;

X_2: objective value given by the user (learner), it is the average of the users' rates and $0 \leq X_2 \leq 1$;

X_3: the number of the key-objects in the video/the number of the key-objects given by the user;

X_4: the period of at least one of the key-objects shown in the video/the length of the video;

X_5: the period of all key-objects shown in the video at the same time/the length of the video;

W_i: the rating weight defined by the system administrator and $\sum_{i=1}^{5} W_i = 100$;

Key-object is the object that contains the keywords and/or RFID tags data given by the user.

(2) Similarity of two videos

After recording a task as video, the user may want to find similar videos to compare. Therefore, the system shows two criteria to help users. The first is the similarity of the objects, including the two videos, is calculated as follows.

$$\sigma = \frac{m}{n}$$

where,

n = the total number of the objects in the two videos,

m = the number of the same objects in the two videos and $0 \leq \sigma \leq 1$.

If all the objects of the two videos are the same, σ becomes 1. If there is no same objects in the two videos, σ becomes 0.

Second, the similarity of the orders of the same objects in the two videos is calculated using Kendall's rank correlation coefficient:

$$\tau = \frac{4P}{n(n-1)} - 1$$

where,

$n =$ the number of the same objects,
$P =$ the sum of the same objects ordered after the given objects by both orders and $-1 \le \tau \le 1$.

If the orders of the objects in the two videos are the same completely, then τ becomes 1. If there is no same order of objects in the two videos, σ becomes 0. The system recommends videos according to the above two criteria.

4. Evaluation

We conducted an evaluation to measure how LORAMS can support ubiquitous learning. The task was cooking fried rice as shown in Figure 6. We choose "fried rice" because it is cooked in a short time, there are a lot of different kinds of cooking methods for this, and the students have cooked it more than once.

4.1. *Design*

Twenty one students from the Department of Computer Science in the University of Tokushima were involved in this experiment for three days. The subjects consist of 8 undergraduates, 8 first-year students in master course, and 5 second-year students in master course. Each of them was given ten minutes to cook.

Before starting the task, we explained the devices to participants and how to use PDA and RFID tag reader. Five tools, eight seasonings and seventeen ingredients such as salt, pepper, egg, oil, pan, chopsticks, rice, onion, green pepper, sausage, soy source and carrot, were prepared and attached to different RFID tags. According to the pre-questionnaire, the students' experiences about cooking were evaluated and they were defined as expert or non-expert. There were eleven students (expert),

Figure 6. Scene of the study.

who had experience in cooking, and there were ten students (non-expert), who had little experience in cooking.

In the first day of study, all subjects cooked fried rice and all their processes were recorded. In the second day, they watched and made annotations into others' videos using LORAMS. In the third day, the non-expert group cooked fried rice again and compared videos using LORAMS.

4.2. *Results*

During the study, 31 videos were registered in LORAMS. The average time for recoding video was 4 minutes and 50 seconds, the maximum was 6 minutes 59 seconds, and the minimum was 2 minutes 55 seconds (Figure 7). Also the average time for encoding and uploading video was 1 minute and 44 seconds, the maximum being 2 minutes 28 seconds, and the minimum was 1 minute. Thus, the time for encoding and uploading was one third of the video recording time for cooking. Therefore, we believe the waiting time is not so long for the users of LORAMS.

As for the result of Q1, it was a comparatively good result. Even the cooking video recorded by the non-professional people was helpful for cooking.

Table 1. Results of questionnaires.

No.	Questionnaire	Ave	SD
Q1	Was it useful to watch a video using LORAMS for cooking?	4.3	0.59
Q2	Was it easy for you to recognize your mistake and difference with other persons using LORAMS?	4.6	0.47
Q3	Was it easy to use the interface for comparing videos?	4.1	0.86
Q4	Is it easy to learn something by comparing videos?	4.1	0.93
Q5	Overall, was it easy to reflect your mistake on the next time using this system?	4.6	0.51
Q6	Overall, do you want to register and share your video using LORAMS?	4.0	1.00
Q7	Overall, do you want to use this system again?	4.4	0.71

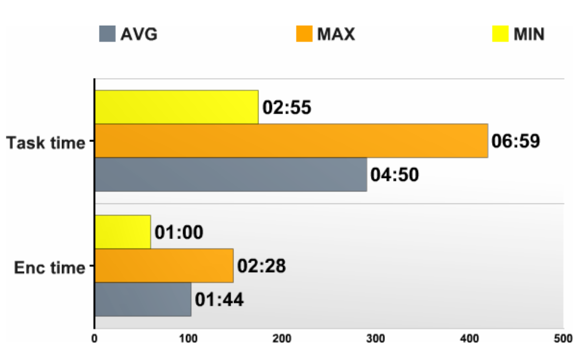

Figure 7. Task and encoding time.

According to the results of Q2 and Q3, the ranking methods for the video recommendation and the similarity were good. Most of the subjects could find the appropriate video for the reference and comparison because the system can easy indentify and compare the objects in the two videos using the RFID tags. The result of Q4 was fairly good. We found that most of the subjects could recognize his/her mistake and distinguish himself/herself by comparing the videos. The subjects found that the difference in objects (17%), timing (41%), actions (42%) are shown in Figure 8. They learnt, for example, that green pepper and carrot should put before the rice, and the rice will be crisp if it is mixed with an egg before putting into a pan, and how to wield a pan.

From Q5, it was easy for most of the subjects to use the interface to compare videos, because we explained how to use it before starting the evaluation. However, it is necessary to improve the interface more. The results of Q6 and Q7 show that most of the subjects could easily find good or bad points for the next time using LORAMS interface. From Q8 and Q9, most of the subjects wanted to share their videos and to use LORAMS again and commented that they are willing to share videos to know how others cook in different ways. However, a few subjects did not want to share their videos because of their failure in cooking.

Next, we focus on a subject and compare the video timeline and objects. In Figure 9, the first timeline is the video by an expert and the second is the subject's. There are two differences between the two timelines. First, the subject did not put any ingredient into a pan as shown at (A). Second, the subject took more than 3 times the period to finish cooking in (B). Therefore, the subject's fried rice was burnt.

The subject understood what was wrong by watching videos and cooked fried rice again. The time of the second cooking is shown in the third timeline in Figure 9. According to the diagram, the cooking was improved by comparing the videos, and

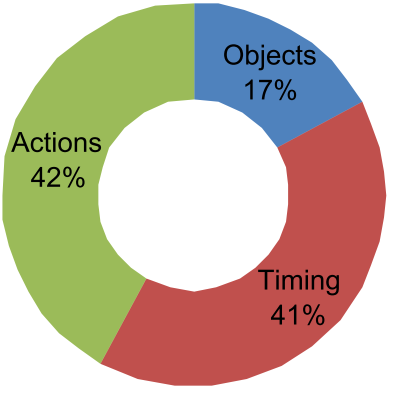

Figure 8. The rate of the difference that the subjects indentified.

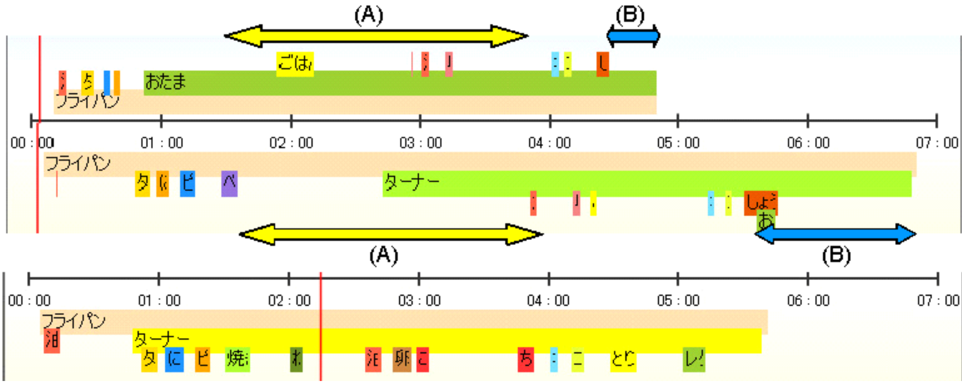

Figure 9. Timelines of three videos.

the two problems mentioned above were solved. Most of the subjects commented that it was easy to find similar or not-similar videos to make comparison, and that the user can easily find the difference between two videos using the graphical timeline. Therefore, LORAMS is a useful tool to share and compare videos. Furthermore, the subjects also learnt that:

(1) They should have taken care about the heat of the gas oven.
(2) All things were fried quickly.
(3) They could use different ingredients.
(4) They should have taken care of the order to put an ingredient into a pan.

5. Conclusions

Ubiquitous computing will be integrated into everyday objects and activities and support not only to provide the right information at the right time at the right place but also to capture, share and reuse human's experiences.

This paper proposes a ubiquitous learning environment called LORAMS (Link of RFID and Movies System), which supports the learners with a system to share and reuse learning experience by linking videos and environmental objects. The system has the following features:

(1) Without any text annotation, the learner can find the suitable scenes that include the objects around the learner.
(2) The system recommends the learners suitable videos for watching and comparing according to the numerical ranking methods that are proposed in this paper.
(3) The system allows the learners to share knowledge through making annotations into videos.

The evaluation was conducted by twenty-one university students and showed the following results:

(1) The students could learn how to cook different kinds of fried rice by watching videos using LORAMS.
(2) Most of the students agreed to the ranking of the search results and the video recommendation for comparisons.
(3) Overall, it was easy to find your mistakes or differences by comparison and to using LORAMS. Also most of them wanted to use the system again.

This evaluation was conducted in a short term study. We will conduct a longer-term experiment. Also we will compare LORMAS with a video sharing system like YouTube in terms of the benefits of RFID-based tagging and the video-comparison function.

In future work, we will improve the user interface and ranking methods based on the students' comments. Also we believe LORAMS can be applied to many domains in which a person needs several kinds of objects and skills to do a task, and the result varies depending on the objects and the skills used in the task; for example, checking upon cars such as oils, battery, and tires, second language learning for the people who are living in a foreign country, or surgery operations and chemical bioreactor experimentations. In those domains, the videotaping and RFID tagging in LORAMS are used purposely to support further teaching or learning, rather than "just tape it and use it in some day". Finally, ubiquitous computing society is still not realized currently, but we believe we should start to design learning environments for the future.

References

Abowd, G. D., & Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing, *ACM Transaction on Computer-Human Interaction 7*(1), 29–58.

Arias, E., Eden, H., Fischer, G., Gorman, A., & Scharff, E., Beyond Access (1999). Informed participation and empowerment. In *Proc. Conference on Computer Supported Collaborative Learning (CSCL '99)*, (pp. 20–32).

Borriello, G. (2005). RFID: Tagging the world, *Communications of the ACM*, *48*(9), 34–37.

Bush, V. (1945). As we may think, *The Atlantic Monthly 176*(1), 101–108. Retrieved July 17, 2008, from http://www.theatlantic.com/doc/194507/bush

Davis, M. (1993). An iconic visual language for video annotation. In *Proc. IEEE symposium on visual language*, (pp. 196–202).

Fischer, G. (2001). User modeling in human–computer interaction, *Journal of User Modeling and User–Adapted Interaction (UMUAI) 11*(1/2), 65–86.

Gemmell, J., Bell, G., & Lueder, R. (2006). MyLifeBits: A personal database for everything, *Communications of the ACM*, *49*(1), 88–95.

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., & Wood, K. R. (2006). SenseCam: A retrospective memory aid. In *Proc. of conference on human factors in computing systems (UbiComp 2006)*, (pp. 177–193).

Hsi, S., & Fait, H. (2005). RFID enhances museum visitors' experiences at the exploratorium, *Communications of the ACM*, *48*(9), 60–65.

Kawamura, T., Fukuhara, T., Takeda, H., Kono, Y., & Koide, M. (2007). Ubiquitous memories: A memory externalization system using physical objects, *Personal and Ubiquitous Computing 11*(4), 287–298.

Lindgren, R., Pea, R., Lewis, S., & Rosen, J. (2007). Learning from digital video: An exploration of how interactions affect outcomes. In *Proc. computer–supported collaborative learning (CSCL2007)* (pp. 444–446). Mahwah, NJ: Erlbaum Associates.

Lyytinen, K., & Yoo, Y. (2002). Issues and challenges in ubiquitous computing, *Communications of ACM 45*(12), 63–65.

Nagao, K., Shirai, Y., & Squire, K. (2001). Semantic annotation and transcoding: Making web content more accessible, *IEEE Multimedia 8*(2), 69–81.

Ogata, H., & Yano, Y. (2004). Context–aware support for computer–supported ubiquitous learning. In *Proc. IEEE international conference on wireless and mobile technologies in education (WMTE2004)*, (pp. 27–34).

Ogata, H., Matsuka, Y., El-Bishouty, M. M. & Yano, Y. (2007). LORAMS: Linking physical objects and videos for ubiquitous learning. In *Proc. international conference on computers in education (ICCE 2007)*, (pp. 297–304).

Sakamura, K., & Koshizuka, N. (2005). Ubiquitous computing technologies for ubiquitous learning. In *Proc. IEEE international workshop on wireless and mobile technologies in education (WMTE2005)*, (pp. 11–18).

Smith, J. R., & Lugeon, B. (2000). A visual annotation tool for multimedia content description. In *Proc. SPIE Photonics East, Internet Multimedia Management Systems, 4210*, 49–59.

Sugimoto, M., Hosoi, K., & Hashizume, H. (2004). Caretta: A system for supporting face-to-face collaboration by integrating personal and shared spaces. In *Proc. conference on human factors in computing systems (CHI2004)*, (pp. 41–48).

Yamamoto, D., & Nagao, K. (2005). Web-based video annotation and its applications, *Journal of the Japanese Society for Artificial Intelligence*, *20*(1), 67–75.