Effects of video reflection by wearable memory enhancement system

Yasuhiro Kamo
Tokyo Metropolitan University
kamo@krmgiks5.tmit.ac.jp

Yasushi Ikei
Tokyo Metropolitan University
ikei@computer.org

Abstract

This paper describes the effects of video reflection provided by a wearable memory aid system that assists the memorization process targeted to the action performed for the use of real-world objects. The goal of the system is to exploit infinite capacity of human memory by assisting memorization using recorded video and other information captured during interaction process to real-world information objects. The enhanced memory recall about the user's action to the object helps the user access physically or virtually in mind vast materials distributed in the real-world for intellectual production works. Memorization reinforcement is achieved by drawing attention, promoting rehearsal by flashing a video of interaction, and organizing memory with relevant information. The action of the user is captured by a cap-mounted miniature camera, an RFID reader, and a posture sensor attached to the user. For video-aided rehearsal, a small HMD (Head Mounted Display) presents the captured scenes immediately after the interaction, and/or at appropriate times for the user's context. We conducted the second basic model experiment in the research program to substantiate our idea. A model task to memorize the page images in a book was performed by the subjects, which suggested the memory enhancement system could reinforce memorization of the user's visual experience by video aided rehearsal.

1. Introduction

Human performs intellectual tasks using a variety of information materials distributed in the real world. To facilitate the task with easy access over the materials, the whole association of the materials needs to be memorized. An approach to augment availability of the materials is to utilize the spatial property of them efficiently in the form of visual memory. In the memory it is stored in the first-person viewpoint and probably augmented by emphasizing the scene with some rendition to change it compact and of impact.

Although the research of context-aware computing is gradually incorporating the user's status to establish intelligent agents to manage distributed information, to exploit further the large human capacity in memorizing information is considered as a right complement of the alternative approach that effectively works in conjunction with the computer agent-based approach.

Machine memory aid has been discussed in various ways among which the episodic memory aid is most relevant to this study. In the Forget-me-not system [1] the user's context of action was automatically recorded to provide the user who forgot past actions with easy retrieval interface that worked on fragments of related items in his/her episodic memory. The location, people around, events, and time stamps constitute the database for the later retrieval. In [2], the proactive presentation of related information based on the context by a wearable remembrance agent was discussed. The goals of their system are not necessarily in strengthening the memory trace in the user's brain itself, but in helping the retrieval of relevant items by the cues easily memorable for human.

In this paper we introduce a wearable memory-aid system that aims to support human memorization of the action of interaction with objects that he/she uses in everyday work environment. The objects in general include books, magazines, bulletin reports, paper materials, stationeries, information devices, appliances, etc. that hold information related to our intellectual activities. An experiment that demonstrates the utility of the system on memorizing page images of a book is discussed.

2. Wearable Memorization Aid

Memorization is a process to encode and store the perceived information with a specific representation in the human brain. During the process, if the elaborative rehearsal is effectively performed, we can usually retrieve the memory item at a high success rate. Before the process, information being encoded needs to be perceived first with an adequate amount of attention that enables a reliable recall.

We propose memorization aid in the two aspects, drawing attention and promoting rehearsal, for enhancing memory retrieval by using a wearable apparatus. The wearable memorization aid system observes the user's behavior to

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

handle objects in the real world, and presents cues in the course of the action. The cues are built primarily based on the captured interaction records such as visual scenes and sounds the user saw and heard when he/she manipulated the objects. The property data of objects and related links stored in the database are also involved in the cues.

Objects handled are labeled and identified by RFID tags that are read by a tag reader the user wears. The interaction is segmented on the tag detections. The cues are presented during the interaction, immediately after the interaction and at any adequate time when the user is in a resting state to prefer reviewing the action. Thus, the system performs detection and recording the user's interaction, and then presents the recorded visual and audio cues to the user.

2.1. Capturing a Video

The user's interaction is detected by wearable and environmental sensors. Changes in interaction state that are important for the cueing are the start of contact of the user to an object, the end of the contact, releasing the object, and the beginning of a resting state. The detection is performed by an RFID system and an arm posture sensor that measures the acceleration indicating the progress of interaction.

During the user's interaction, the scene and the sound are captured by a camera attached to the brim of a cap that covers the workspace area around the arm's reach. The camera view involves an object, a part of the user's hands and the environment showing a scene observed from the first person's viewpoint.

2.2. Video Flashback

Cueing messages of the system are presented in the following three phases. a) At state changes, for drawing attention of the user to the change occurred in the interaction, b) During the interaction, for providing association to the object, and c) After the interaction, for aiding rehearsal at appropriate timing with adequate quantity.

The presentation procedures are as follows. a) State changes are informed by any of a short sound, a screen blink, a color change, or tactile cues. b) Relevant information such as the title of related objects and events as well as the target (held) object are presented on a wearable display while the user holds the object. c) A video that recorded the process of interaction is shortened and retouched, then played as flashback on a wearable display for appropriate times immediately after the interaction is terminated.

The benefits of this memorization aid are expected to:

- Help build a correct index map (cognitive map) of information media in the real environment,
- Recapture the interaction, or reestablish memory trace in case attention was insufficiently paid to that interaction,

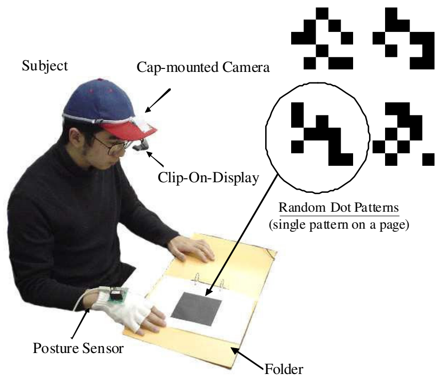

Figure 1. Experimental apparatus and a pattern to memorize

- Facilitate a later search within the captured interaction data with added cues, which reduces unwanted extended interaction to the database,
- Intensify consciousness of engagement by projecting a part of the user's figure in the captured video, and
- Organize the memory of interaction with existing consistent schema of memory by providing relevant data to augment the association of the interaction.

Problems expected and to be addressed are as follows.

- Presented cues may cause a distraction when user's attention is directed to a different task.
- Repetition of cueing may increase a cognitive load of the user.
- Similar images of interaction may decrease the uniqueness of each cue for recall.
- Context behind the user's action needs to be estimated accurately to optimize timing and frequency for appropriate presentation.

2.3. Hardware Settings

A subject wore a cap that implemented a small CMOS image sensor camera (Kanebo Ltd.) with a wide-angle lens (120-degree) for capturing a video of the user's visual field while performing a task. A video was stored in QVGA (320x240) format at 15 frame/s. An RFID tag reader (V700 Series, Omron Corp.) was attached to the user's palm to detect an object when it was held. A small VGA display (CO-3, MicroOptical Corp.) was fixed to the bow of glasses to show a video. A computer was placed nearby on the table to relieve the load to the subject during the experiment.

3. Video Reflexion over Paper Browsing

Information provided by printed materials is one of principal targets for our intellectual activities. We examined application of the memory enhancement system to browsing behavior of paper materials. The object to memorize

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

is the images of pages in books or documents. The images seen from the user's viewpoint is flash-backed again immediately after browsing pages. Since a page image itself is static, it is expected that reflective presentation needs not necessarily dynamic but a static (freeze frame) video would be sufficient for the rehearsal. Thus, a sequence of static frames extracted from page images captured during browsing is presented. Two forms of video presentations were used to compare different allocation of playback duration to each frame. The uniform-still-video presentation (U-mode) has the same duration for each page. The specific-still-video presentation (S-mode) gives different duration for each page.

The item to memorize in the experiment is a random dot pattern as shown in Figure 1. The pattern spans 12 x 12 cm^2 with 5 x 5-matrix elements in which ten elements are filled black. The random dot pattern was printed at the center of the page of an A4-size (210 x 297 mm^2) sheet. A sequence of seven patterns randomly selected from 90 patterns produced on a computer was put in a folder. Two blank pages at the first and the last were added to the pages to mask the item for controlling the exposure time.

3.1. Procedure

The two video presentation modes, the uniform(U)-still and the specific(S)-still videos, and the additional none(O) mode in which no video was played back were randomized and balanced to eliminate the order effect. The subject performed the task three times for each mode in total. A session consisted of three runs of different modes. The runs are separated by a five-minute rest.

The subject was instructed to flip over the seven pages on a signal given by the experimenter every five second. The subject tries to memorize the seven patterns. After the observation of the last page, the freeze frame video showing the seven pages from the first page is played back during which the subject performed rehearsal with the images.

In the U-mode, every page image is equally presented for three-second duration. The S-mode presented the seven page images with different durations shown in Table 1. This allocation of time is intended for canceling the serial position effect in which the first and the last pages (items) and their neighbors are well recalled. A longer duration in the middle is for the expected lower recall performance. After this video playback, a one-minute delay was inserted before the subject was asked to recognize the seven patterns from a list which drew 30 patterns on a sheet. No video presentation (O) mode also involved the one-minute blank before starting recognition. Then the subject took a five-minute rest before the next run.

Table 1. Duration of each page image in the U- and S-modes [sec]

Page	1	2	3	4	5	6	7
U-mode	3.0	3.0	3.0	3.0	3.0	3.0	3.0
S-mode	1.5	2.8	4.0	4.5	4.0	2.8	1.5

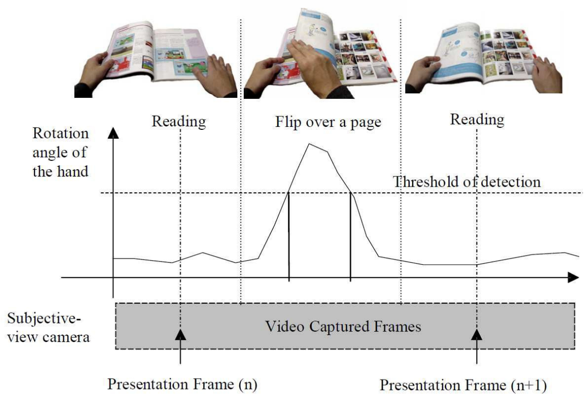

Figure 2. Detection of flip-over and presentation frame

3.2. Apparatus and subjects

A 3D posture sensor was attached to the back of the subject's hand. The compass in the sensor detected hand motion during the flip-over of pages as shown in Figure 2. The presentation frame was captured when hand motion stopped and the subject was expected to be gazing at the pattern on the page. This detection procedure is obviously not intended for general purpose but appropriate for this experiment. Six subjects of the mean age of 24.3 from the institute participated in the experiment.

3.3. Result

The scores of the sessions and average are depicted in Figure 3. A perfect (correct) recognition will score 70 (10 points for each pattern). With the U-mode and S-mode video playback the subject could recognize the pattern significantly better ($p<.01$) than the no video mode. Obviously the difference between U-mode and S-mode is not significant. It seems that the both presentations equally assisted the rehearsal of pattern images. Figure 4 shows the recognition performance at each page. Regarding the O-mode and U-mode, the score did not indicate the serial position effect of increased performance at the first and the last tasks; the recognition performance was relatively constant with a slight decreasing trend. The extinction of recency effect may be attributed to the one-minute blank before starting recognition. In the S-mode the performance in the middle pages looks rather increased inversely to the usual serial position effect. This should be accounted for the increased duration of frames presented for the pages in the middle.

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

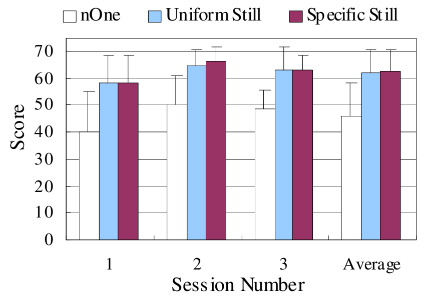

Figure 3. Recognition performance (in sessions and average)

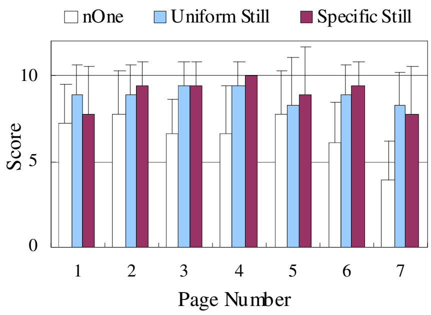

Figure 4. Recognition performance (in pages)

4. Discussion

The wearable memory enhancement system detects the user's interaction to information media objects in the real workspace and presents a digest video of the interaction seen from the user's viewpoint. The video playback seems to have aided rehearsal efficiently on how the user manipulated the media. This visual promotion of memory is a good way to make an intense clue to the media since it reinforces episodic memory of the user. The episodic memory is more susceptible to forgetting than semantic memory in general.

People usually try to recall their action as episodic memory partly based on their visual memory; however it tends to be unclear since they often do not pay sufficient attention to such outward property of the media as its geometry or location while attending its content. However, the physical aspect of media should actively be projected to human memory as well as its contents, since the surface aspect, location and appearance, are considered to increase essential association independent of its content. The video flashback is expected to add a vivid and accurate reality in memory trace through this efficient rehearsal. Most subjects reported that the video itself was quite helpful since their action, even though consciously and spontaneously performed, vanished rapidly in mind was recaptured by the video recorded at a first-person viewpoint. The results of the experiment showed that memorization was promoted by flash-backing to an eye of the scene where object shifting and figure browsing were performed.

The experiment of paper browsing exhibited the effectiveness of the video flashback. The recognition performances were almost the same with the U-mode and the S-mode. However the subjective evaluation obtained by a questionnaire indicated that the U-mode was easier to perform rehearsal than the S-mode. This result was supported by the answer to another question for time arrangement. The reason was the inappropriate arrangement of replay duration; too large difference between the middle page and the first/last pages. We need further experiment on the freeze frame presentation.

5. Conclusion and Future Work

We developed the wearable memory enhancement system and demonstrated effectiveness of video-aided rehearsal on the model task. The system looks potentially promising in various application areas that involve intellectual production, education and rehabilitation. However the system is still a prototype at the first stage and has many limitations. First, we need to elucidate the principle of video aided memorization more intensively to be able to answer what kind of video clip is effective for rehearsal or what sort of video image is easy to remember. Not very much is known about video memorability that would be controlled on modern video editing technology.

Second, the system devices need more flexible control options to enable versatile applicability to human actions and the objects. The presentation design should permit more flexible timing, caption on the video, and multi-modality displaying. All these need the finer context sensing which is a main problem of wearable computing as well. Practically more important issue is the timing control of presentation in the course of the user's activity. To determine when to flashback, or not to do, based on the contextual importance of objects is a challenging aspect of this research. Utility of the wearable memory enhance system depends much on this decision.

Acknowledgement

The authors would like to thank Mr. Yoji Hirose at Sony Ericsson Mobile Communications Japan, Inc. for his considerable contribution of the experiment.

References

[1] M. Lamming and M. Flynn. Forget-me-not: Intimate computing in support of human memory. In *FRIEND21: Int. Sympo. on Next Generation Human Interface*, pages 125–128, 1994. 1

[2] R. J. Rhodes. The wearable remembrance agent: a system for augmented memory. In *ISWC'97*, pages 123–128, 1997. 1