# A Lifelog Viewer System Supporting Multiple Memory Cues

Jiaming Zhang[(✉)], Jie Liang[(✉)], and Jiro Tanaka[(✉)]

Graduate School of IPS, Waseda University, Fukuoka, Japan
zhangjiaming777@toki.waseda.jp, liangjie@fuji.waseda.jp,
jiro@aoni.waseda.jp

**Abstract.** Lifelog means using wearable devices to record human's life automatically which can help users recall. More and more lifelog cameras with multiple sensors become available to the public. To view the vast amount of lifelog photos captured by the devices, the manufacturers provide supporting viewer systems. However, current systems only enable users to view the photos in time order by installing the specialized software. Previous research stated that memories can be recalled by memory cues, such as where the event happened, who was involved. In this research, we intend to find out the most important memory cues that help people recall and create a new lifelog viewer system to help users recall past more efficiently without increase the burden on users. To find out which cues appear more in events remembered by people, we conducted several experiments. The proposed new lifelog viewer system can recognize object, location, face and time information in lifelog photos automatically and show photos according to these memory cues information instead of in time order. With the system, users can retrieve photos of past events via these multiple memory cues. We also conducted a preliminary user study to evaluate, the results show the validity of the proposed system.

**Keywords:** Lifelog · Autographer · Memory recall · Memory cues combination · Computer vision

## 1 Introduction

The term "lifelogging" appeared in the 1980s. Steve Mann did experiments with a wearable personal imaging system designed and built by himself to capture continuous physiological data [1]. Since then, more and more work has been conducted in this field. Progressive miniaturization by the computer industry has enabled wearable devices to become less obtrusive. In the 2000s, lifelog cameras with multiple sensors entered the market and became available to the public such as SenseCam, Vicon Revue and Autographer [2]. These cameras can be used for memory aids by capturing digital record of wearer's life, including photos and other sensor data. Manufacturers also provide supporting viewer systems for reviewing this information. For example, SenseCam provides a PC-based viewer application that can manage and replay image sequences and sensor data captured by the device [3]. The application allows users play an image

sequence at different speed. Users can delete individual images from the sequence if they are badly framed or of inferior quality. Also, users can create bookmarks with certain images which can be used to navigate the image in a long image sequence. Autographer also provide a desktop software for users to view photos and select photos to share [5]. Each photo is along with a GPS coordinate, and the additional captured metadata, such as ambient light, temperature, etc.

However, these systems only enable users to view the photos in time order. Some previous research state that people usually think of several elements in the event when they recall past moments [7]. For example, who was involved, what they used or saw, when and where the event happened. Time is not the only cue helps people recall. We call these elements that can help people recall as memory cues.

With embedded sensors and GPS unit in lifelog devices, we can get information like the positions where the photos were taken. Moreover, there are many available computer vision services [11, 12] which make it possible to extract useful information from lifelog photos automatically without users' effort. This potential information could be used as memory cues to help users recall past.

In this paper, we propose a new lifelog viewer system to help user recall past more efficiently without increase the burden on users. The system can extract memory cues information automatically by integrating with several services. With the viewer system, users can retrieve related photos of past events via multiple memory cues.

## 2   Goal and Approach

In this paper, we present a new lifelog viewer system to help users recall past more efficiently. Before designing the system, we should verify whether lifelog can help users recall more events and find out the most important memory cues in helping people recall.

To achieve these goals, we conduct experiments which consist of record part and recall part. Based on the experiment results, we develop the new lifelog viewer system. The system can recognize important memory cues information in lifelog photos automatically and enable users to retrieve related photos of past events via specific memory cues.

### 2.1   Experiment Setup

**Device.** We use a wearable lifelog camera called Autographer as our experiment record device. Autographer is a hands-free, wearable device which can take photos automatically every 30 s.

Autographer has 6 built-in sensors, including GPS positioning, accelerator, light sensor, magnetometer, motion detector and thermometer to help determine when to shoot [5]. For example, when the wearer runs through the bus, the camera's accelerator will automatically sense the movement and take a photo or when the wearer goes from a warm bar to the snowy street to greet friends, the camera will also take a photo automatically by sensing the change of temperature. There's an official desktop software of Autographer for users viewing photos in time order.

**Experiment Sheet Design.** Each sheet represents an event. Participants should write down each event with a brief description and four memory cues including *who*, *what*, *when* and *where*. *Who* means the subjects related to this event, if the participant is the only person in this event, he/she can write "myself". *What* means the specific objects participant interacts with. *When* means the occurred time of the event, which is accurate to hour. *Where* means the specific place where the event happens. At last, participants should write down the memory cues which he/she thinks are most important to help him/her recall this event.

## 2.2  Procedure

We invited 12 participants (6 males and 6 females) ranging in age from 22 to 28. We gave participants a brief explanation about the experiment and told them how to use Autographer, then let participants wear Autographer to collect their one-day life for continuous 6 h between 9am to 7pm. And participants can switch off the device during their private time.

**Experiment 1.** The first experiment is the record part. Immediately after 6-h recording, all lifelog photos were imported into the Autographer viewer system. We divided all participants into two groups. Participants in Group 1 wrote down the events they can remember by themselves not using any reference (Type A) at first, and then viewed the lifelog photos in Autographer viewer system to write down some additional events (Type B). Participants in Group 2 viewed lifelog photos with Autographer viewer system directly and wrote down all the events (Type C).

**Experiment 2.** The second experiment is the recall part. We gathered all participants and asked them to recall events again after a month. The division of group and event types are the same with the first experiment. For example, participants in Group 1 first recalled what happened one month ago by themselves and viewed the previous recorded photos to write down more events.

## 2.3  Results

We collected 196 (63 (Type A) +34 (Type B) +99 (Type C)) sheets in the first experiment and 158 (27 (Type A) +47 (Type B) +84 (Type C)) sheets in the second experiment. The sum of Type A and B was almost the same with the number of Type C, which means that independent thinking before using lifelog doesn't make sense. The total average number of recalled events decreased because some memories of what happened one month ago were missing. The events recalled by independent thinking decreased a lot (Type A), from 10.5 to 4.5. However, participants could recall more events with the help of lifelog (Type B and Type C), the total average amount only dropped from 16.5 to 14 (see Fig. 1). It indicated that lifelog is indeed useful in helping people recall past.
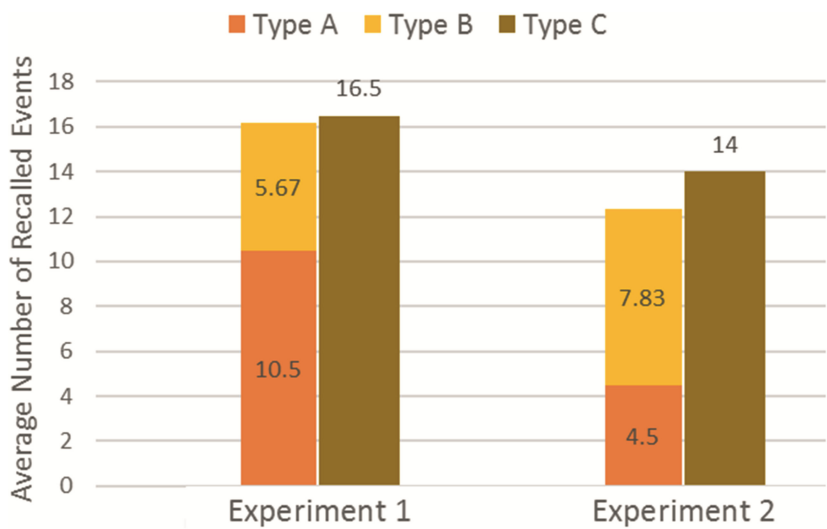
**Fig. 1.** Comparison of the average number of recalled events between experiment 1 and experiment 2.



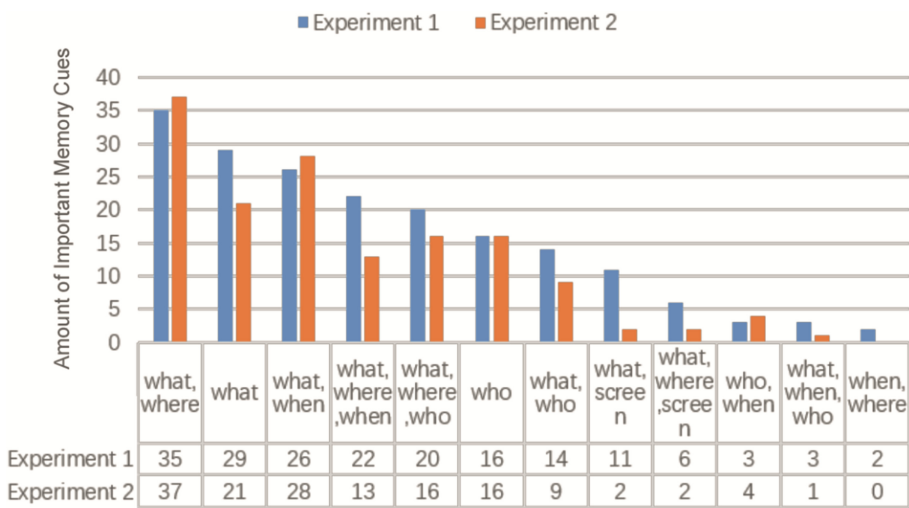| | what, where | what | what, when | what, where ,when | what, where ,who | who | what, who | what, screen | what, where ,screen | who, when | what, when, who | when, where |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Experiment 1 | 35 | 29 | 26 | 22 | 20 | 16 | 14 | 11 | 6 | 3 | 3 | 2 |
| Experiment 2 | 37 | 21 | 28 | 13 | 16 | 16 | 9 | 2 | 2 | 4 | 1 | 0 |

**Fig. 2.** Comparison of memory cues result between experiment 1 and experiment 2.

We also counted the important memory cues written by participants. Unexpectedly, *when* and *who* were not as important as we imagined. Instead, *what* was the most important as a single cue, followed by *who*. The result showed that participants preferred to write down multiple memory cues. And despite the results slightly changed, the rankings had no significant difference between these two experiments: the combination of *what* and *where* ranked first. Although it turned out that *when* was almost useless as a single

cue, but it could be useful when combined with other memory cues, such as *what* and *when* (see Fig. 2).

In summary, we can say that current viewer systems only enable users to view lifelog photos in time order are not good enough in helping users recall.

## 3   Lifelog Viewer System

We propose a web-based lifelog viewer system that can extract object, location, face and time information from lifelog photos automatically by integrating with Microsoft cognitive services [12, 13] and Google Maps API [14]. These kinds of information are corresponding to the important memory cues we found through our experiments, including *what*, *where*, *who* and *when*. Our proposed system enables users to retrieve photos of events via these multiple memory cues, instead of only viewing photos in time order, which helps users recall past events more efficiently.

### 3.1   Scenarios

The user met one of his/her friends and wanted to recall what happened with this friend in the past. He/she could view the photos he imported to our system and double click the avatar image of this friend to view the related photos contains the face of this friend. Or the user remembered he/she went to a convenient store but forgot what he/she bought. He/she could double click the location image of the store to view the related photos.

Another suitable scenario is, the user remembered he/she read some books at a bookstore, but he/she forgot what the title of a book was and wanted to recall the scene. He/she could click cue images of book and the bookstore to query related photos. Then the system would show the result photos captured at the bookstore, which contain books after user clicked the query button.

### 3.2   System Design

The system mainly provides users the following four functions to enable users to import lifelog photos captured by Autographer, view the photos in categories of memory cues that we found important through the experiments and retrieve photos via cues to recall past events.

The lifelog photos import function, which requires user to import photos captured by Autographer to the system.

The memory cues extraction function, which will automatically recognize the memory cues information in lifelog photos, including object, location, face, and time. Then the system will show the cue images in these four categories (see Fig. 3).
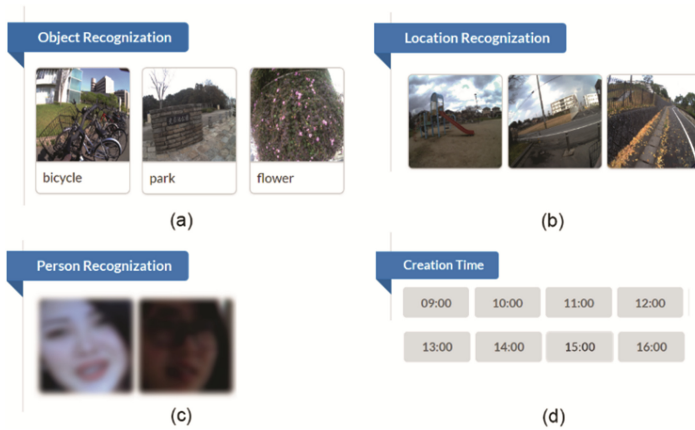
**Fig. 3.** Cue images in four categories. Object (a), Location (b), Person (c), Time (d).

The view photos via single memory cue function, which enables user to view related photos with one specific cue by double clicking the cue image. Because the results of experiments indicated *when* is useless as a single cue, so time cannot be used independently.

The view photos via cues combination function, which enables user to retrieve photos via multiple memory cues. If there are no photos satisfied the query, the system will send a feedback and inform user to choose another cue.

**Process of Recalling Events with Proposed System.** User needs to plug the Autographer into his/her computer and select lifelog photos to import to the system by accessing the website of the system at first (see Fig. 4). After system extract memory cues information, user can view and retrieve photos via cues.



**Fig. 4.** Import lifelog photos.

*View Photos via Single Cue.* User double clicks the object cue image. The retrieved photo sequence is organized in time order. Each photo retrieved is tagged with the date it was taken and the location. User can view these photos by clicking next or previous button (see Fig. 5).
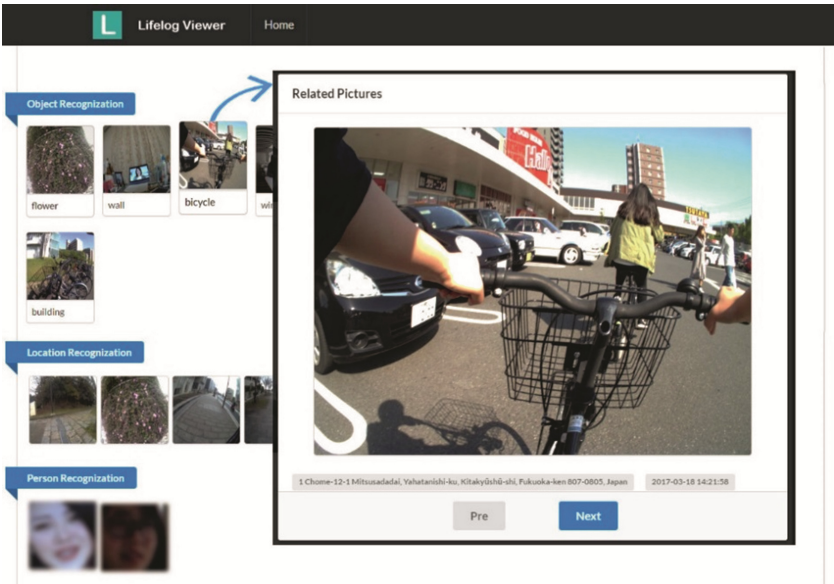
**Fig. 5.** View photos sequence retrieved by single memory cue.

*View Photos via Multiple Cues.* User can click cue image to select it as a query cue, all the chosen ones will be shown on the right of the page, user also can deselect a query cue by clicking it. After determining the cues to be selected, user can click the query button to view related photos (see Fig. 6).
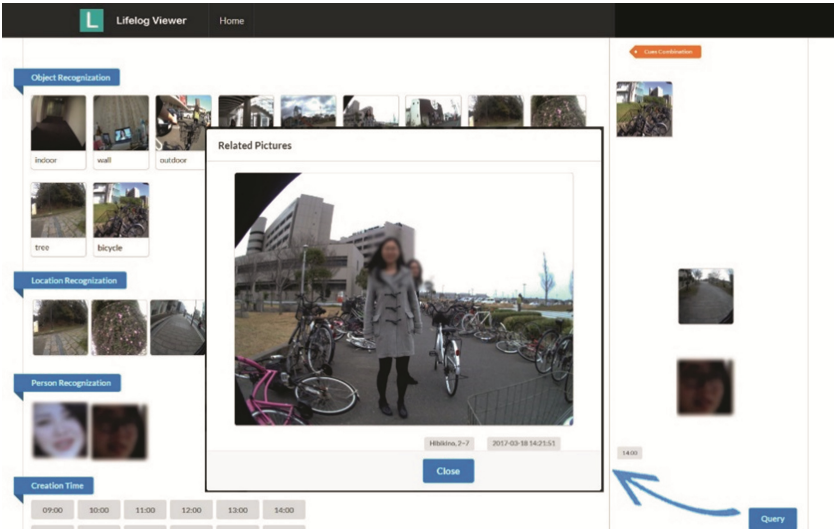


**Fig. 6.** View photos sequence retrieved via multiple cues.

### 3.3 Implementation

The system is implemented based on Browser/Server pattern and mainly uses the combination of the Spring Boot [13] and the Hibernate [11]. The system design follows the principles of the MVC. View is responsible for rendering output results in webpages, which are coded using HTML, CSS and JavaScript. Controller is for handling requests, which is coded in Java. Model is responsible for managing the data [10]. We make use of a MySQL database to store data. The database mainly contains picture, avatar, object, location entities, which are used to store photos in byte arrays and related memory cues information.

User uses our system by accessing the website via browser, each manipulation will send a request to our server, the request will achieve to the Spring Dispatcher Servlet and the Spring Boot will dispatch the request to related services. There are four services in our system, including object service, location service, avatar service and time service, which are integrated with the services from Microsoft and Google to handle the object, location, face and time information in photos. The integrated services include Computer Vision API, Face Recognize API, and Google Maps API. The services will connect with database and deal with data via Hibernate Framework and return the retrieved data to the browser. A MySQL database is used to store the data of memory cues information extracted from the photos (see Fig. 7).
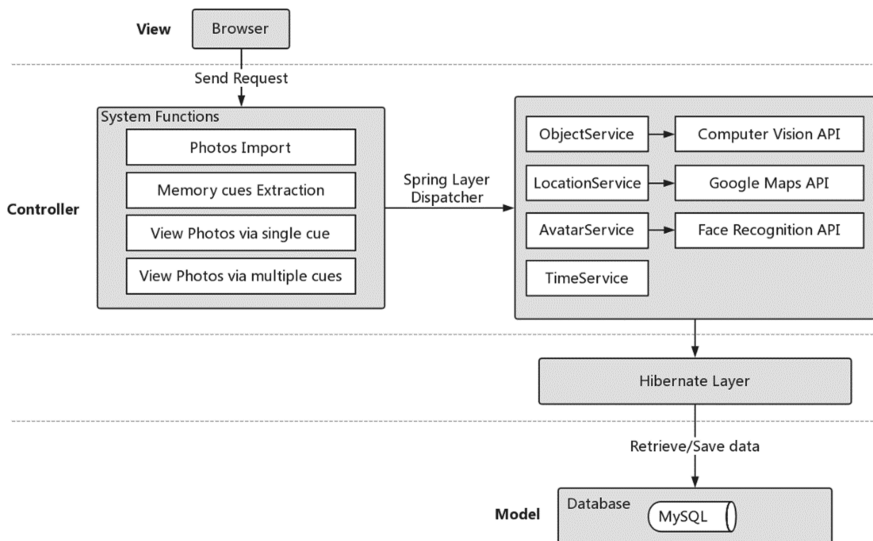


**Fig. 7.** System architecture.

**Object Service.** This service handles all requests related to objects, including recognize objects from photos, store object information into database, show object cue images and query photos via objects.

Microsoft Computer Vision API [12] is used to recognize objects in photos. This API can return a list of object tags by uploading the image bytes of the photo and sending

the request to http://westus.api.cognitive.microsoft.com/vision/1.0/analyze, each object in the returned list has its name and confidence. The system will only store objects with confidence over 0.8. This threshold of confidence is set to filter unreliable results.

**Location Service.** This service is responsible for all requests related to location, including extract location information from photos, store location information into database, show location cue image and query photos via location.

Google Maps API [14] is used to deal with the location information in photos. The system uses this API as a dependency with Maven. For each photo, the system will load latitude and longitude from image bytes and get address information by calling the function GeocodingApi.reverseGeocode().

**Avatar Service.** This service handles all requests related to face, including recognize faces from photos, store face avatar into database, show avatar cue images and query photos via the face of person. Microsoft Face API [13] is used to detect faces in photos.

To detect faces and store the avatar for each person's face, the system will create a face list with a specified ID at first to store the faces recognized from photos by sending request to http://westus.api.cognitive.microsoft.com/face/v1.0/facelists. Then, the system will process each imported photo to detect faces by uploading the image bytes of the photo and sending request to https://westus.api.cognitive.microsoft.com/face/v1.0/detect. Each detected face has a unique ID and the rectangle area of the face in the photo.

The detected faces will be added into the face list the system maintains by sending request to http://westus.api.cognitive.microsoft.com/face/v1.0/facelists/persistedfaces. The required parameters contain the ID of the face list which is created before and the rectangle area of the face in the lifelog photo. Then the system will find similar faces with the new detected face in the face list by sending request to https://westus.api.cognitive.microsoft.com/face/v1.0/findsimilars. If the returned similarity confidence is over 0.6, the face will be considered as a new, unique face. An avatar will be created for each unique detected face.

**Time Service.** This service handles request related to time information, including extract capture time of photos and query photos via time. Photos captured by Autographer have a specific naming format, the system extract the time information from the filename of photos and store it in a specific format. The hour of the captured time is also stored in units of half an hour.

## 4 Evaluation

We conducted a preliminary user study to verify whether our proposed system is more efficient to help users recall than current viewer system.

**Procedure.** We invited 4 participants (1 males and 3 females) ranging in age from 22 to 25. All participants were asked to capture photos with Autographer for 3 h. After 3 days, we gathered all participants and divided them into two groups. Participants were

asked to import photos to the lifelog viewer system. Specifically, participants in Group 1 used Autographer viewer system, and participants in Group 2 used our proposed system. Then we let participants finish several tasks (see Table 1).

**Table 1.** Tasks description

| Task | Description |
|------|-------------|
| 1 | Given a single cue, like *what*, retrieve related photos and recall the event |
| 2 | Given multiple cues, like *what* and *where*, retrieve related photos and recall the event |
| 3 | Recall all events as much as possible |

**Results.** We collected a total of 1535 lifelog photos from 4 participants, average photos for per participant is 383.7.

Task 1 is used to investigate whether using single memory cue like *what*, *where* or *who* to retrieve photos is more efficient than viewing photos in time order. Task 2 is used to investigate whether using multiple memory cues like *what* and *where*, or *who* and *where* to retrieve photos is more efficient than viewing photos in time order. Participants in Group 1 took 12.5 s for completing Task 1 and 21.5 s for Task2 on average. Participants in Group 2 took 7.5 s for completing Task 1 and 9 s for Task2 on average. The results show that users can retrieve photos of past events in a shorter time with our proposed system. In particularly, when given multiple memory cues, using our proposed system is much more efficient (see Fig. 8).
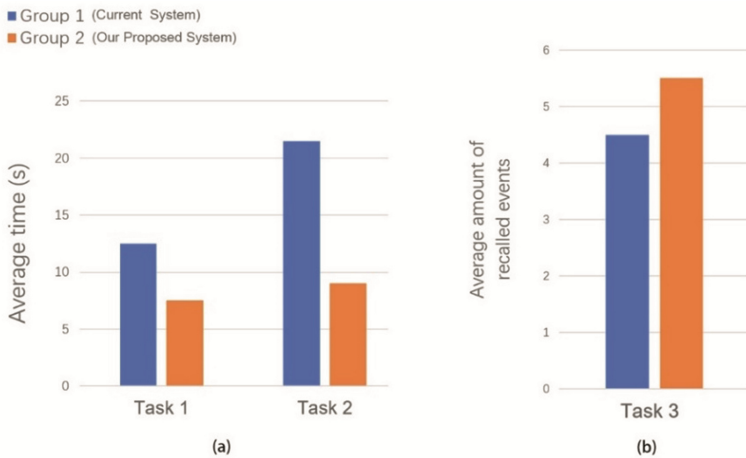


**Fig. 8.** Tasks results. Comparison of average time of completing Task 1 and Task 2 (a). Comparison of average amount of recalled events in Task 3 (b).

Task 3 is used to verify whether our system can help users recall. The average amount of recalled events of participants in Group 1 is 4.5 and participants in Group 2 recalled 5.5 events on average. The result shows that our proposed system can help users recall past events as current Autographer system does (see Fig. 8).

## 5   Related Work

### 5.1   Memory Cue

Episodic memory [6] enables individuals to remember their personally experienced past. This kind of memory can be recalled by memory cues. Wagenaar spent 6 years on his study of his autobiographical memory from 1978 [7]. He recorded all events by means of four aspects, what the event was, who was involved, and where and when it happened. In the recall phase he cued himself with one or more of these four aspects to recall the other aspects. He found *what* as a single cue was very powerful, and *when* was almost useless. He also stated combination of cues was more effective than single cues.

### 5.2   Lifelog System

With lifelogging technology, we can capture everything we experienced. Sellen et al. found evidence that SenseCam images do facilitate people's ability to recollect past events [4]. With the vast number of lifelog photos, how to extract valid information and retrieve events more conveniently are nowadays researchers focusing on.

Memon developed a prototype system using an Android smart phone that can recognize the present situation and search for relevant past episodes [9]. In his research, there are three key elements: people in sight, objects of interaction and present location. With the system, lifelog related to current location or the specific location which users defined can be accessed directly. Likewise, lifelogs can be retrieved based on people and objects currently present near the user.

Matsuoka designed a system by linking lifelog photos with tags [8]. In her system, users need to manually select photos and edit tags of cues information one by one at first. Users can enter keywords to retrieve events and the system will show related photos and suggest related tags as feedback. This series of operations increase the burden on the users and not friendly enough.

Comparing with above mentioned systems, our work has several advantages. First, our system can automatically recognize object, location and face information from lifelog photos by integrating with services like Microsoft Computer Vision API and Google Maps API, which relieves users' burden. While Matsuoka's system requires users to manually select photos and enter tag information. Second, we make use of Autographer, which can capture photos automatically and unobtrusively. The prototype device in Memon's system requires users to capture manually and needs several other alternatives, like RFID sensors for individuals' identification.

## 6   Conclusion and Future Work

In this paper, we present a new lifelog viewer system for helping users recall past events more efficiently. By integrating with services like Computer Vision API and Google Maps API, the system can extract object, location, face and time information automatically. These kinds of information are corresponding to the important memory cues we

found through the experiments, including *what*, *where*, *who* and *when*. Users can retrieve photos via multiple cues to recollect the past events. The results of the preliminary user study indicate that user can retrieve lifelog photos and recall past events more efficiently with our proposed system.

In the future work, we will improve the system further. For example, the processing time of the vast amounts of photos is a problem, which affects the user experience. And Matsuoka [8] stated that *screen* could be a potential memory cue because people spend considerable time on computing devices. *Screen* means the display of computing devices users interact with. Our system may support *screen* cue by collecting computer and smartphone usage data of users in the future studies.

## References

1. Mann, S.: Wearable computing: a first step toward personal imaging. Computer **30**(2), 25–32 (1997)
2. Zhou, L.M., Gurrin, C.: A survey on life logging data capturing. In: SenseCam Symposium (2012)
3. Hodges, S., et al.: SenseCam: a retrospective memory aid. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 177–193. Springer, Heidelberg (2006). https://doi.org/10.1007/11853565_11
4. Sellen, A., Fogg, A., Aitken, M., Hodges, S., Rother, C., Wood, K.: Do life-logging technologies support memory for the past? An experimental study using SenseCam. In: CHI 2007 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 81–90. ACM, California (2007)
5. Thoring, K.C., Mueller, R. M., Badke-Schaub, P.: Ethnographic design research with wearable cameras. In: CHI EA 2015 Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 2049–2054. ACM (2015)
6. Tulving, E., Donaldson, W.: Organization of Memory. Academic Press, New York (1972)
7. Wagenaar, W.A.: My memory: a study of autobiographical memory over six years. Cogn. Psychol. **18**, 225–252 (1986)
8. Matsuoka, R.: Understanding how cues in lifelog affect memory recalls. University of Tsukuba (2016)
9. Memon, M., Mahoto, N., Khuhawar, F., Tanaka, J.: Retrieval of life logs based on users context. Sindh Univ. Res. J. (Sci. Ser.) **47**(4), 851–860 (2015)
10. Gupta, P., Govil, M.C.: Spring Web MVC Framework for rapid open source J2EE application development: a case study. Int. J. Eng. Sci. Technol. **2**, 1684–1689 (2010)
11. Xia, C.-L., Yu, G.-C., Tang, M.: Efficient implement of ORM (Object/Relational Mapping) use in J2EE framework: hibernate. In: International Conference on Computational Intelligence and Software Engineering, pp. 1–3 (2009)
12. Cognitive services Computer Vision API reference. https://azure.microsoft.com/en-gb/services/cognitive-services/computer-vision/. Accessed 26 Jan 2018
13. Cognitive services Face API reference. https://azure.microsoft.com/en-gb/services/cogntive-services/face/. Accessed 26 Jan 2018
14. Google Maps API reference. https://developers.google.com/maps/. Accessed 26 Jan 2018