Proposal of a Memory Support System Utilizing First-Person Perspective Snapshots for Older Adults

1st Keiko Kodama
Graduate School of Engineering
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
kodakoma@es4.eedept.kobe-u.ac.jp

2nd Takuya Nakata
Center of Mathematical and Data Science
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
tnakata@bear.kobe-u.ac.jp

3rd Sinan Chen
Center of Mathematical and Data Science
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
chensinan@gold.kobe-u.ac.jp

4th Sachio Saiki
School of Data and Innovation
Kochi University of Technology
185 Miyanokuchi, Tosayamada, Kami, Kochi, 782-8502, Japan
saiki.sachio@kochi-tech.ac.jp

5th Masahide Nakamura
Center of Mathematical and Data Science
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
masa-n@cmds.kobe-u.ac.jp

6th Kiyoshi Yasuda
Graduate School of Engineering
Kobe University
1-1 Rokkodai-cho, Nada, Kobe, 657-8501, Japan
yasukiyo.12@outlook.jp

***Abstract*—In Japan, the prevalence of Mild Cognitive Impairment (MCI), a prodromal stage of dementia, has been increasing with population aging, becoming a significant social issue. MCI impairs memory, particularly episodic memory, imposing psychological and time-related burdens on both individuals and caregivers. Although several methods have been proposed to support memory by collecting and presenting daily life data, many rely on third-person perspectives or require complex device operations. Therefore, it is difficult for such individuals to operate existing systems independently. The goal of this paper is to support older adults with MCI in independently collecting and storing first-person perspective data, thereby facilitating self-initiated recall of episodic memories. Our key idea is to integrate a wearable device built with a Raspberry Pi and cloud storage. We propose a system that automatically captures first-person view images, stores them in the cloud, and enables search based on tags assigned through object recognition. This system aims to support self-initiated recall of episodic memories by allowing older adults to independently collect and retrieve contextual data from their own perspective.**

***Index Terms*—episodic memory, declining birthrate and aging population, memory impairment, MCI, cloud Storage**

I. Introduction

The global rise in aging populations has become a significant social issue, necessitating multifaceted forms of support [1]. In particular, Japan is experiencing rapid population aging, with individuals aged 65 and over comprising 29.1% of the total population in 2023—the highest proportion in the world [2]. Consequently, the prevalence of dementia and Mild Cognitive Impairment (MCI), a prodromal stage of dementia, is increasing [3]. MCI is characterized by noticeable declines in cognitive function and memory, reported by the individuals themselves or their families, while basic daily activities remain intact [4]. It especially affects recent episodic memory, leading to difficulties such as forgetting appointments, medication schedules, or the locations of personal items. Episodic memory refers to the recollection of personally experienced events, and by nature, it is difficult even for close family members to compensate for its loss. As a result, older adults with MCI often experience anxiety and frustration in daily life without knowing how to cope. In addition, their caregivers and family members are frequently burdened with repeated questions and unpredictable behavior, creating both psychological and time-related stress [5]. For these reasons, there is a growing need for systems that allow older adults with MCI to independently compensate for episodic memory loss.

Various systems have been proposed for collecting and presenting daily life data to support memory. However, many of these systems rely on third-person perspectives, such as stationary monitoring devices, and fail to capture the subjective experience of the individual. Moreover, most systems require users to manually sift through large volumes of data or add annotations, which poses usability challenges for older adults unfamiliar with digital devices. These barriers make it difficult for such individuals to operate existing systems independently, without assistance from others.

Specifically, three key challenges remain unresolved:

- **Challenge 1 (C1): Older adults are unable to collect first-person perspective data.**

- **Challenge 2 (C2): Older adults are unable to independently collect and store such data.**
- **Challenge 3 (C3): Older adults are unable to independently search for data that serves as cues for episodic memory recall.**

To address these challenges, this study proposes a system that automatically captures first-person view images, stores them in the cloud, and enables search based on tags assigned through object recognition. This paper particularly focuses on addressing C1 and C2, and the proposed system is composed of the following five approaches:

- **A1: Wearable device with camera module**
  A small camera module is worn around the neck to automatically capture images from the user ' s visual perspective at regular intervals.
- **A2: Capturing snapshots and adding timestamps**
  Each captured image is automatically annotated with the date and time it was taken, enabling time-based recall.
- **A3: Uploading to the cloud**
  Annotated images are automatically uploaded to the cloud via Wi-Fi, eliminating the need for manual storage management.
- **A4: Automatic tagging of images**
  The uploaded images are analyzed using object recognition, and relevant tags are assigned automatically for improved searchability.
- **A5: Retrieval and display of cues for episodic memory recall**
  Users can retrieve images by searching with time or tag-based queries, allowing efficient and self-reliant recall of past events.

This system aims to support self-initiated recall of episodic memories by enabling older adults to independently collect and retrieve contextual data from their own perspective.

Based on the above approaches, we developed a wearable device equipped with a camera module. Additionally, as a case study, we examined whether the system could successfully perform automated image capture and storage, and explored methods for automatic tagging.

II. Preliminaries

A. Current Status of MCI

The aging of the population is recognized as a global issue, requiring multifaceted forms of support [1]. In particular, Japan is experiencing significant population aging, with individuals aged 65 and over comprising 29.1% of the total population in 2023, the highest proportion in the world [2]. Along with this demographic shift, the prevalence of Mild Cognitive Impairment (MCI) has been increasing year by year, becoming a serious social concern [3].

MCI is considered a prodromal stage of dementia, in which basic daily activities remain intact, but the individual or their family notices declines in cognitive function and memory [4]. Episodic memory, especially recent events, is particularly vulnerable in MCI, often leading to difficulties such as forgetting medication, meals, or the location of objects and appointments. Since episodic memory refers to personally experienced events, it is inherently difficult even for close family members to compensate for. As a result, older adults with MCI frequently experience anxiety or frustration in daily life due to memory lapses. Moreover, memory impairments caused by MCI affect not only the individual but also those around them, often exceeding the capacity of informal support. People with MCI may repeatedly ask the same questions or exhibit unexpected behavior, placing a psychological and temporal burden on caregivers and family members [5]. In addition, due to trends such as the decline of multi-generational households, the proportion of older adults living alone is expected to increase, reaching 26.1% for men and 29.3% for women by 2050 in Japan [2]. As a result, it is becoming increasingly difficult for older adults to receive support from those around them. For these reasons, systems are needed that enable older adults with MCI to compensate for episodic memory loss independently, without relying on external support.

B. Related Work

1) Challenges in Data Collection: To support episodic memory recall, the use of first-person perspective data—captured by the individual who experienced the event—is considered effective. Many studies have investigated the recording and storage of daily life data to support memory, using modalities such as images, videos, audio, and accelerometer data [6]. Among these, approaches based on perspective snapshots have shown both high acceptance among people with MCI and promising results in aiding episodic memory recall [7]. However, much of the research on data collection for older adults has focused on monitoring or anomaly detection, using third-person data such as fixed cameras [8]. This leads to a lack of data that reflects the subjective experience of the individual. Although some methods for collecting first-person data have been proposed, they often have limitations such as capturing only held objects [9] or requiring prior registration of target objects or locations. Such requirements make these systems difficult to use independently, especially for older adults unfamiliar with digital devices. While it is theoretically possible to manually record daily events using smartphones or written memos, many older adults lack the habit of doing so, due to a cognitive bias that underestimates their own memory decline.

Thus, realizing a system that allows older adults to independently collect first-person perspective data remains a challenge.

2) Challenges in Data Storage and Retrieval: Because the volume of collected data is typically large, technologies that support efficient storage and retrieval of relevant information are essential. With the growing availability of wearable devices, some commercial products now enable automatic logging of first-person perspective data. However, many still require manual uploading to the internet or connection to a PC to access the data. In addition, most existing systems simply display the collected data, requiring users to manually

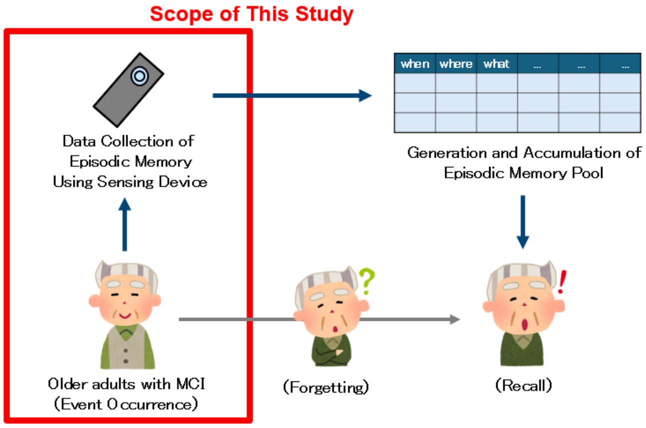


Fig. 1: The overall architecture of the proposed system

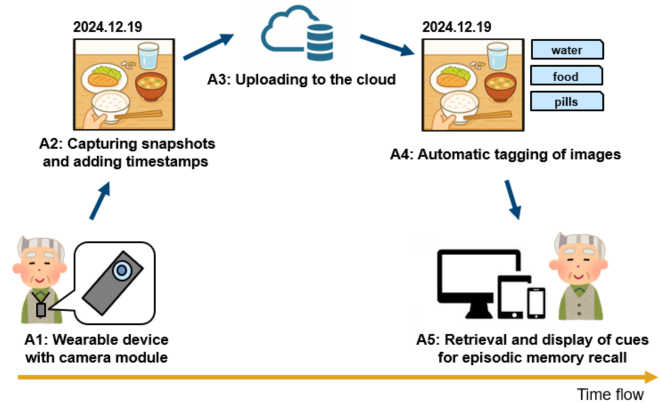


Fig. 2: The system architecture in the scope of this study

search for the relevant information themselves [9] [10]. While systems that allow retrieval of data relevant to episodic memory recall do exist, many require manual tagging in advance, making them difficult for older adults to use independently.

Therefore, realizing a system that enables older adults to independently store and retrieve first-person perspective data, and to use it for episodic memory recall, remains a challenge.

C. Identified Challenges

The following three issues remain open challenges in enabling older adults with memory impairments to independently compensate for episodic memory loss.

- Challenge 1 (C1): Older adults are unable to collect first-person perspective data.
- Challenge 2 (C2): Older adults are unable to independently collect and store such data.
- Challenge 3 (C3): Older adults are unable to independently search for data that serves as cues for episodic memory recall.

The overall architecture of the proposed system is shown in Fig. 1. In this system, perspective snapshots are automatically captured at regular intervals using a wearable camera placed near the user ' s neck, allowing the images to reflect the user ' s natural visual experience. These images are then tagged using object recognition and stored in a searchable format. By enabling older adults to retrieve past information as needed, the system aims to address the three challenges listed above. This paper specifically focuses on addressing C1 and C2.

III. Proposed Method

A. Goal and Key Idea

The goal of this study is to support older adults with MCI in independently collecting and storing first-person perspective data, thereby facilitating self-initiated recall of episodic memories. The key idea is to integrate a wearable device built with a Raspberry Pi and cloud storage. To achieve this, we propose a system that periodically captures perspective snapshots from a neck-worn camera and stores them in the cloud. These images are then automatically tagged using object recognition, enabling efficient image retrieval. Specifically, the system addresses C1 (Older adults are unable to collect first-person perspective data) and C2 (Older adults are unable to independently collect and store such data) through the following five components:

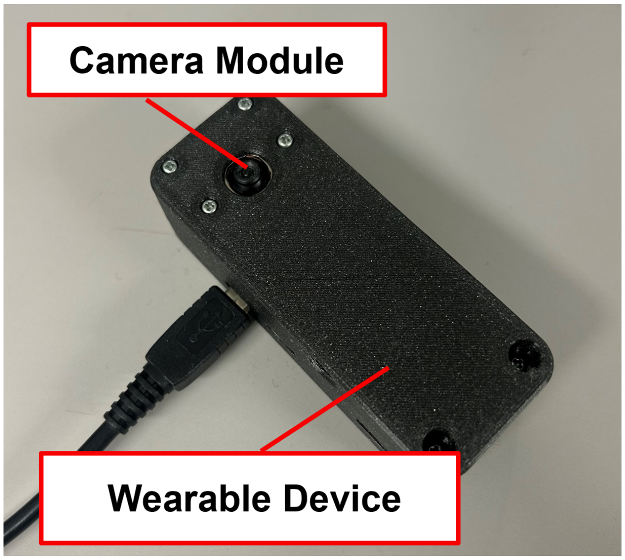
- A1: Wearable device with camera module
- A2: Capturing snapshots and adding timestamps
- A3: Uploading to the cloud
- A4: Automatic tagging of images
- A5: Retrieval and display of cues for episodic memory recall

B. System Architecture

The system architecture in the scope of this study is shown in Fig. 2. The system consists of a wearable device with a camera module, cloud storage, a user, and an interface for displaying search results. The user wears the device around the neck while using the system. The wearable device automatically captures perspective snapshots at regular intervals, appends metadata at the time of capture, and uploads the data to the cloud. This enables older adults to collect first-person perspective data without needing special operations or prior knowledge, thereby addressing C1 and C2. Additionally, the stored images are automatically tagged using object recognition, and relevant images are displayed when the user searches using tags.

C. A1: Wearable Device with Camera Module

The wearable device serves as the foundation for data collection in the proposed system. It is equipped with a camera module and is designed to be worn around the user ' s neck, providing a first-person perspective data. This approach allows the system to capture images closely aligned with the user ' s natural field of vision, which is critical for creating meaningful episodic memory cues. To operate, the device automatically captures perspective snapshots at regular intervals. By utilizing this device, older adults can accumulate a continuous visual



(a) The appearance of the implemented device

(b) A perspective snapshot taken using the proposed system

Fig. 3: The details in this implementation

record of their daily activities without requiring deliberate actions or technical knowledge.

D. A2: Capturing Snapshots and Adding Timestamps

The snapshots obtained in A1 are automatically annotated with the time they were taken. This information enhances the searchability of the data by allowing users to filter and retrieve images based on when they were captured. As a result, the system generates a continuous set of structured image records, such as "image file + timestamp", at regular intervals. This structured dataset supports episodic memory recall by enabling chronological search patterns that naturally align with how human memories are typically recalled.

E. A3: Uploading to the Cloud

The images annotated with timestamp information in A2 are automatically uploaded to the cloud via Wi-Fi. As users do not need to manage physical storage or perform manual operations, the system is accessible to older adults with limited digital literacy. Storing the data in the cloud also ensures that large volumes of real-time information can be preserved securely, without being limited by the storage capacity of the wearable device.

F. A4: Automatic Tagging of Images

The images saved to the cloud in A3 are subsequently analyzed using object recognition to assign relevant tags. Each image is analyzed to detect visible objects or features, and relevant keywords are assigned based on the results. Given the large volume of accumulated image data, this is crucial for efficient retrieval. By automating the tagging process, the system eliminates the need for users to manually label or sort the images, thereby reducing both cognitive and operational burden. This process eliminates the need for manual annotation, allowing users to accumulate their visual data in a searchable format without added burden.

G. A5: Retrieval and Display of Cues for Episodic Memory Recall

When users want to revisit past events, they can perform searches using the timestamps annotated in A2 and the tags assigned in A4. By combining temporal and semantic metadata, the system allows users to efficiently narrow down large volumes of image data without having to browse through each image manually. This reduces the need for manual browsing and supports a self-reliant search experience. The retrieved images are then displayed through an interface designed for easy browsing, helping older adults recall their episodic memories more easily and with greater confidence.

IV. Implementation

A. Technologies Used

The appearance of the implemented device is shown in Fig. 3(a). We used the Raspberry Pi Zero 2 W, developed by the Raspberry Pi Foundation [11], as the wearable device, and the Raspberry Pi Camera Module 3 as the camera module [12]. The following technologies were used:

- Programming language: Python
- Services: Nextcloud, PhotoPrism

Nextcloud [13] is an open-source cloud storage platform that enables secure and flexible file storage and sharing. Because it has no storage limit, it can manage large volumes of image data. In the proposed system, it is used for managing perspective snapshots. PhotoPrism [14] is an open-source photo management application equipped with object recognition capabilities using Google TensorFlow. It automatically tags

Fig. 4: Image of a remote control used for tagging evaluation

image data and enables tag-based search. In our system, it is used for tagging and searching perspective snapshots.

V. Case Study

As a case study, the proposed system was used to capture perspective snapshots at regular intervals and automatically annotate them with timestamps. Additionally, we evaluated the automatic tag generation and search functionality using PhotoPrism.

A. Capturing and Storing Perspective Snapshots

Fig. 3(b) shows a perspective snapshot taken using the proposed system. As shown in Fig. 3(b), a clear image from a first-person perspective was automatically captured, with the timestamp added in the upper right corner of the image. These images were uploaded to Nextcloud via Wi-Fi, enabling time-series-based retrieval for episodic memory support.

B. Tagging and Searching Perspective Snapshots

To evaluate the automatic tagging and search capabilities of PhotoPrism, we prepared an image in which a remote control is clearly visible, uploaded it to PhotoPrism, and examined the tags it generated. Fig. 4 shows the image used in this evaluation.

The automatically generated tags for Fig. 4 were “electronics” and “brown”. Similar results were obtained with other images. For example, an image showing a person’s face clearly was tagged only with “portrait” and “brown”. Likewise, an image of the sky received only the tag “blue”, and an image showing many trees was tagged only as “brown”. These results indicate that the tags generated by PhotoPrism were highly limited, typically reflecting dominant colors or broad categories rather than specific objects. Therefore, the tagging was insufficiently accurate for object-based image search and is not considered suitable for supporting episodic memory recall.

VI. Discussion

A. Tagging Method for Perspective Snapshots

In this study, PhotoPrism was used as the method for tagging and retrieving perspective snapshots. Although this tool offers automatic object recognition, the resulting tags were limited in both accuracy and expressiveness. Even when we provided clear images of target objects such as a remote control the assigned tags were often too generic, including labels such as “electronics” or “brown,” and failed to convey content that would be meaningful for recalling personal experiences. In this system, accurate tagging plays a vital role in allowing users to search through a large collection of images to find cues for memory retrieval. It is especially important that the tagging process captures not only the types of objects but also contextually important details that may help users recall specific events. Therefore, PhotoPrism is considered unsuitable as a tagging method for perspective snapshots in this system.

Moreover, the evaluation in this study was conducted using still images taken under stable conditions. However, in actual use, the wearable camera will take pictures while the user is moving, which may cause images to be blurred or partially obstructed. These conditions are expected to make the tagging task even more difficult, increasing the need for reliable and precise recognition.

In the future, it will be necessary to explore more advanced object recognition technologies and generative AI-based tagging methods to enable more effective tagging. For example, services such as Google Cloud Vision API could be introduced to enhance object detection and expand the variety of available labels. Additionally, recent image recognition models that integrate visual and language understanding may help generate more detailed and context-aware descriptions, capturing not only individual objects but also entire scenes and activities. These improvements could contribute to providing a more effective and supportive experience for memory recall.

B. Privacy Protection

The proposed system continuously captures perspective snapshots from the user’s environment, which raises potential privacy concerns. For example, documents or mail placed on a desk may contain private information such as addresses or phone numbers, which could be unintentionally recorded. Additionally, the constant awareness of being recorded may cause psychological stress for both users and those around them, potentially discouraging long-term use of the system. Moreover, since the camera is worn on the body and captures surroundings without active input, users may find it difficult to maintain full awareness of what is being recorded at any given time. To address these concerns, the system should be improved to better accommodate privacy protection. Possible enhancements include implementing a masking function when personal information is detected, such as blurring text or faces, and providing users with an intuitive way to manually toggle the system on and off as needed. Incorporating indicators or alerts to make users aware of active recording status may also help mitigate psychological discomfort and enhance transparency in system use.

VII. Conclusion

This study proposed a system that supports older adults with MCI in independently collecting and storing first-person data to assist with self-directed recall of episodic memories. The system captures perspective snapshots, stores them in the cloud, and enables image retrieval by automatically tagging them using object recognition technology. Specifically, a wearable device using a camera module was developed. It was confirmed that the system can automatically capture perspective snapshots at regular intervals, append contextual information, and store them in the cloud. In addition, the tagging method for the stored images was evaluated.

Future work includes implementing more advanced object recognition techniques to improve tagging effectiveness. Additionally, we aim to enhance the system practicality by addressing privacy concerns and ensuring stable operation during everyday activities.

Acknowledgment

This research was partially supported by JSPS KAKENHI Grant Numbers JP25H01167, JP25K02946, JP24K02765, JP24K02774, JP23K17006, JP23K28091, JP23K28383.

References

[1] World Health Organization: Ageing and Health: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health (Visited on 2025-4-18).

[2] Cabinet Office, Government of Japan: Annual Report on the Ageing Society: (2024).

[3] Kyushu University: Study on the Prevalence and Future Estimation of Dementia and Mild Cognitive Impairment in Japan, pp. 9–12 (2024).

[4] R C Petersen and J C Morris: Mild Cognitive Impairment as a Clinical Entity and Treatment Target, Archives of Neurology, vol. 62, pp. 1160–1163 (2005).

[5] T Maruo, Y Ito, N Katakura, M Uda, N Ogawara: Everyday Challenges of People Diagnosed with Mild Cognitive Impairment and their Family Caregivers, Kobe City College of Nursing, vol. 26, pp. 57–65 (2022).

[6] V Sunil, N Marmasse, C Schmandt, T Martino: An Audio-Based Personal Memory Aid, Ubicomp 2004, LNCS, vol. 3205, pp. 400–417 (2004).

[7] O Gelonch, M Ribera, N Codern-Bové, S Ramos, M Quintana, G Chico, N Cerulla, P Lafarga, P Radeva, M Garolera: Acceptability of a Lifelogging Wearable Camera in Older Adults with Mild Cognitive Impairment: A Mixed-Method Study, BMC Geriatrics, vol. 19, Article 110 (2019).

[8] K Takahashi, D Yokoyama: Remembrance Support, Journal of the Robotics Society of Japan, vol. 28, no. 9, pp. 1082–1083 (2010) (In Japanese).

[9] T Yagi, T Nishiyasu, K Kawasaki, M Matsuki, H Sato: Go-Finder: A Registration-Free Wearable System for Assisting Users in Finding Lost Hand-Held Objects, IPSJ Interaction 2021 Proceedings, pp. 739–744 (Mar. 2021).

[10] Narrative Clip2: https://getnarrative.com/ (Visited on 2025-03-31).

[11] Raspberry Pi Zero 2W: https://www.raspberrypi.com/products/raspberry-pi-zero-2-w/ (Visited on 2024-12-19).

[12] Raspberry Pi Camera Module 3: https://www.raspberrypi.com/products/camera-module-3/ (Visited on 2024-12-19).

[13] Nextcloud: https://nextcloud.com/ (Visited on 2024-12-19).

[14] PhotoPrism: https://www.photoprism.app/ (Visited on 2024-12-19).