VIMES : A Wearable Memory Assistance System for Automatic Information Retrieval

Bermejo, Carlos

ACM
2020-10-12

Bermejo , C , Braud , T , Yang , J , Mirjafari , S , Shi , B , Xiao , Y & Hui , P 2020 , VIMES : A Wearable Memory Assistance System for Automatic Information Retrieval . in MM '20 : Proceedings of the 28th ACM International Conference on Multimedia . ACM , New York, NY , pp. 3191-3200 , ACM International Conference on Multimedia , United States , 12/10/2020 . https://doi.org/10.1145/

http://hdl.handle.net/10138/352318
https://doi.org/10.1145/3394171.3413663

unspecified
acceptedVersion

VIMES: A Wearable Memory Assistance System for Automatic Information Retrieval

Carlos Bermejo
cbf@cse.ust.hk
HKUST

Tristan Braud
braudt@ust.hk
HKUST

Ji Yang
jyangaa@cse.ust.hk
HKUST

Shayan Mirjafari
shayan@cs.dartmouth.edu
Dartmouth College

Bowen Shi
bshi@cse.ust.hk
HKUST

Yu Xiao
yu.xiao@aalto.fi
Aalto University

Pan Hui
panhui@cse.ust.hk
HKUST and University of Helsinki

ABSTRACT

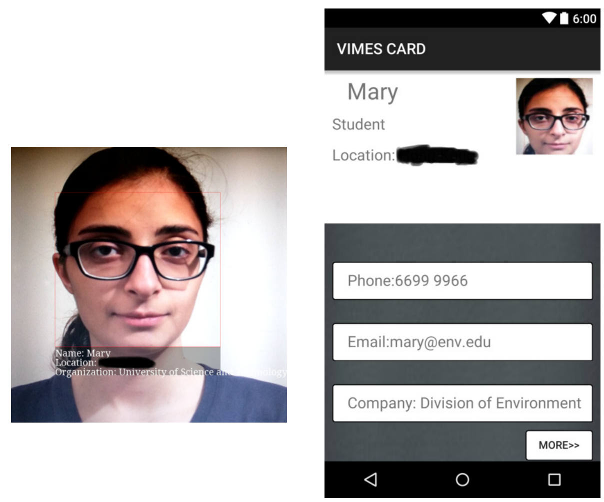
The advancement of artificial intelligence and wearable computing triggers the radical innovation of cognitive applications. In this work, we propose VIMES, an augmented reality-based memory assistance system that helps recall declarative memory, such as whom the user meets and what they chat. Through a collaborative method with 20 participants, we design VIMES, a system that runs on smartglasses, takes the first-person audio and video as input, and extracts personal profiles and event information to display on the embedded display or a smartphone. We perform an extensive evaluation with 50 participants to show the effectiveness of VIMES for memory recall. VIMES outperforms (90% memory accuracy) other traditional methods such as self-recall (34%) while offering the best memory experience (Vividness, Coherence, and Visual Perspective all score over 4/5). The user study results show that most participants find VIMES useful (3.75/5) and easy to use (3.46/5).

KEYWORDS

Augmented reality, wearable, video, face recognition, event detection, memory assistance

ACM Reference Format:

Carlos Bermejo, Tristan Braud, Ji Yang, Shayan Mirjafari, Bowen Shi, Yu Xiao, and Pan Hui. 2020. VIMES: A Wearable Memory Assistance System for Automatic Information Retrieval. In *Proceedings of the 28th ACM International Conference on Multimedia (MM '20), October 12–16, 2020, Seattle, WA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3394171.3413663



(a) Interface on smartglasses. **(b) Interface on smartphone.**

Figure 1: VIMES in action. The system automatically detects the faces of people and extracts information from the recorded interaction. Users can then visualize a summary of their previous information on the smartglasses' screen (a) or on the phone (b).

1 INTRODUCTION

Millions of people are affected by various forms of memory problems [60]. *Dementia* covers various diseases characterized by a cognitive decline that considerably affect people's abilities and daily life. It involves diverse symptoms, including short-term memory loss, having difficulty recognizing relatives and friends, or becoming unaware of the time and place [14]. Even healthy human beings often forget essential information in their daily activities. There is thus undoubtedly a need for memory aids, such as reminder notes [4, 35], for dementia patients and healthy people alike.

With recent developments in wearable computing and multimedia processing, wearable devices can provide real-time memory assistance through automatic information retrieval from the embedded sensors. People often forget everyday details and events, and solutions based on summarizing techniques can considerably improve memory recall [36]. Multiple empirical shreds of evidence show the potential of wearable cameras as a viable solution to remediate autobiographical memory impairment [2, 17, 20]. Several works propose wearable cameras with recognition features such as the face and objects, to improve users' memory recall [17, 20]. The rise of these 'lifelogging' wearables, such as smartglasses [23], can impact the adoption, such as the rising concern of privacy [37, 45]. Moreover, individuals' opinions about these wearable devices' ability to collect data differ, depending on whether they are bystanders

ACM ISBN 978-1-4503-7988-5/20/10…$15.00
https://doi.org/10.1145/3394171.3413663

or users. Further understanding of the constraints and concerns is necessary to design these devices. While these approaches have brought exciting insights, they lack user involvement in the design of the device and further analysis of the possible privacy and adoption challenges of these 'lifelogging' devices.

In this work, we present *VIMES*, a comprehensive wearable memory assistance system for helping users to remember daily-life facts events (also known as also known as *declarative memory*). *Remembering acquaintances* is a typical use case of declarative memory. We consider the following scenario: *A user meets a new person for the first time. When meeting this person a second time, days or weeks later, the user hardly recalls their names or other information about them.* Addressing this use-case is highly dependent on the users' own preferences regarding lifelogging and its applications. Thus, we co-design VIMES with a panel of 20 participants to define the requirements, functionalities, aesthetics, and interactions for a wearable memory assistance system for daily use. Following the insights of the co-design study, we develop *VIMES* as an augmented reality (AR) system on smartglasses for supporting declarative memory in daily life. The proposed system extract information from face-to-face conversation using face recognition, speech recognition, and natural language processing techniques, and concisely displays them in augmented reality, as shown in Figure 1. We evaluate the system from both technical specifications and memory recall performance. Our results show that VIMES is efficient in information retrieval (M=0.93) against other memory recall methods. The adoption of the system features such as information recording (e.g., video) depends on the role of the participant (user or bystander). To the best of our knowledge, this is the first comprehensive study that encompasses the entire design process from the early co-design study to the final user evaluation, going through the development of an integrated system for memory assistance.

The **contributions** of this paper are as follows:

- We conduct a collaborative design study (20 participants) of a wearable memory assistance system.
- We develop VIMES around the expectation and requirements defined in the co-design study. VIMES is an AR-based wearable autobiographical memory assistance system that extracts personal profiles and event information from the user's visual and auditive point of view.
- We evaluate the system performance through an extensive evaluation. VIMES operates with high performance within the boundaries of a normal conversation (visual and audio accuracy > 80%, response delay < $1s$).
- We study user performance, as well as their acceptance, privacy, expected social conformity, and adoption intention [44] of VIMES through a 50-participants study. VIMES achieves better memory recall performance (93%) against traditional methods (34%), and provides a better memory experience.

The reminder of this paper is organized as follows. We summarize the most relevant studies related to memory aid in Section 2 before describing our co-design methodology and the system requirements in Section 3. We describe the resulting architecture in Section 4. We then present the system performance evaluation in Section 5 and user evaluation 6. Finally, we discuss the implications of *VIMES* and the results of our final evaluation in Section 7.

2 BACKGROUND AND RELATED WORK

In this section, we first review the state-of-art of memory augmentation techniques and then discuss the wearable assistance systems and the mechanisms used for wearable offloading.

2.1 Cues of Memory Recall

The details and events of everyday experience are often not retained in memory, and solutions based on summarizing approaches can improve the memory recall of individuals [36]. Visual lifelogging [10, 16] captures real-life images through the camera embedded in wearable devices to create a personal photo or video-based memory prosthetics [33]. The goal of lifelogging is to support people's self-awareness and self-management for memory recall [47]. Current research study how to summarize such data and visualize it to reflect on meaningful personal events [25, 34, 48]. SenseCam [21] is one of the first feasible lifelogging cameras. A neck-work sensor-enhanced camera records images and context passively. Several studies have shown SenseCam's benefits for supporting the recall of episodic memories [6, 20, 50]. Other types of cues have been investigated to assist memory recall. Examples include geo-locational cues [28], audio cues [9], such as ambient sound recording, or emotions [27]. Although recording all-time may seem an excellent approach to help users in their daily lives, the amount of resulting information can hinder their autobiographical memory performance due to the lack of a summarizing system. In [61], authors propose a storytelling system to relive users' moments using summarized stories created from photos. The system helps users to remember previous events in a kinematic manner and offers better performance than other state-of-the-art systems. We confirm these findings in our co-design experiment, where most participants required these summarizing systems to cope with the information recorded by the wearable devices.

2.2 Wearable Assistance

A large number of applications are emerging for wearable assistance like emotion sensing [24] and gait recognition [52]. Many of the applications focus on assisting the physically impaired. For example, OpenGlass [57], Chroma [55] and Gear Face Recognition [7] are designed to help visually impaired people. Bachlin *et al.* [3] presents a wearable assistance application for the Parkinsons' disease patients with the freezing of gait symptoms (a sudden and transient inability to move). As mentioned before, SenseCam [21] takes images from a wearable digital camera based on *lifelogging* [50]. It can help people to recall episodic memory by reviewing the images of previous events. Though valuable, it does not provide real-time memory assistance, and therefore cannot allow instant reminding as VIMES does. Authors in [17] take one step further and works as a real-time wearable assistant for cognitive decline. The contributions of these works focus on the design of the system architecture that efficiently incorporates various cognitive engines like face recognition, object recognition, and optical character recognition. As mentioned earlier, emotions are an essential cue for memory recall [27]. Modern wearables include electroencephalogram (EEG) sensors to record users' mental activities such as emotions. *Lifeloggin* applications can also collaboratively collect multiple sources of information from different sources, aggregate these different types of data in cloud-based services, and provide users with richer information about a particular logged event [1]. For example, IoT sensors can collect temperature, humidity, and number of people in a particular room, where the user is logging a particular event.

Our work focuses on inventing a novel method for helping people retrieve cognitive information, such as personal profiles and event information. Furthermore, our work places much emphasis on user experience and intended adoption in addition to the technology.

3 CO-DESIGN STUDY

Collaborative design studies empower users to be part of the design process [26]. Related works on memory aid systems only focus on system design (e.g., computer vision modules [7, 17], summarization techniques [36, 61]). We adopt a co-design approach [46] in which researchers, designers, and potential users share their ideas and cooperate creatively to generate a new wearable memory assistance system.

3.1 Participants

We recruited 20 participants (8 female, 12 male, age 18 to 24: 8, 25 to 34: 6, 35 to 44: 6) around the university campus. Profession demographics (9 computer science students, 2 post-doc civil engineers, 4 environmental engineers, 2 journalist, 2 office managers, 1 post-doc electronic department) and technologies usage: smartphone (1 to 2 hours: 13.3%, 2 to 3 hours: 40%, 3 to 4 hours: 20%, more than 4 hours: 26.7%), and all participants work with computer (2 to 3 hours: 20%, 3 to 4 hours: 26.7%, more than 4 hours: 53.3%). Only three participants wear smart-watches. We focus on young early adopters of wearables, as according to [12], 31% of U.S. population aged between 25-35 years use wearable accessories (2017).

We informed participants that data would be deidentified, and all recorded data will be password-protected and deleted after the study ends. Participants provided informed consent to participate and be audio-recorded. We carried out the study following the local IRB regulations. We rewarded participants with sweets and soft drinks after the completion of the experiment.

3.2 Study Goal

We ask participants to design a wearable memory assistance system collaboratively.

Design space. Description of the design space: *'The system will need to help users to remember daily life events such as meeting people, buying groceries with straightforward interactions and minimal interruptions. The system should be implementable in a wearable device that users can wear.'*

System capabilities. We analyze participants' responses for different characteristics of the system: (i) aesthetics (location of the wearable, size, weight), (ii) functionality (e.g., interface design), and (iii) interaction with the device (e.g., input, feedback).

Event-cues. Details, information to remember about an event.

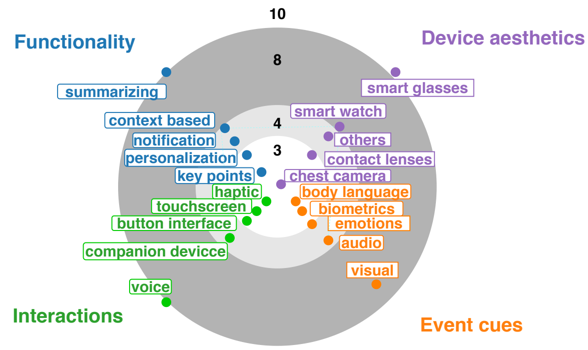


Figure 2: VIMES co-design space. Each circle corresponds to the number of participants that propose the same key idea.

3.3 Protocol

A person with a software and design background helps participants describe their ideas using schemes and photos of different wearable devices, sensors, and locations where wearables can be worn, to help participants during the co-design study. The study is divided into five phases, during which we audio record the participants' answers.

(1) Initial Interview (5 minutes). Participants answer demographic and general technology usage questions.

(2) One-to-one design (18 minutes): We describe the structure of the project as follows: *'We want to design a wearable memory assistance system for autobiographical memory.'* We also describe to participants the wearable capabilities and our design space.

(3) System design (20 minutes): In this part, we gather participants in groups of three to four, where they discuss their design ideas materialized in the previous part. During this phase, we show participants different commercials and prototypes of wearable devices. We describe different familiar places where these wearable devices can be worn, the different capabilities.

(4) Event-Cues (15 minutes): In this phase, we ask each group of participants what they would like to remember from a particular scenario or encounter with other people. Participants propose different use case scenarios: meeting a new classmate on the first day of a course; diary; note-taking application, which can be used for taking class notes. For each proposed situation, we ask participants what details are essential to remember.

(5) Final co-design (15 minutes): In this part, participants summarize their ideas for each of the system capabilities.

3.4 Results

To analyze audio data from the *custom design* study, we iteratively develop a set of key ideas from the interview transcripts [19]. First, two independent researchers read the transcript and developed an initial set of key ideas. Then, both researchers meet in person to discuss and consolidate the key-ideas from participants' transcripts. The inter-coder reliability was 0.73 (Cohen's Kappa). Our goal is to have a breadth of common key-ideas, such as aesthetic designs, event-cues, interactions, and functionality. The two researchers agree on the top key ideas for the prototype design. Figure 2 shows the resulting co-design space for the system capabilities, device aesthetics, and event-cues that the system is required to collect/retrieve. The farther the key ideas are from the center, the more critical they were to the co-design study participants.

Scenarios. Participants propose several use-case scenarios where the *wearable memory assistance system* can help them. P2 said: '*The device can be beneficial during work meetings or lectures as a notes recorder, the system will also summarize the key points of the meeting.*' Other participants (P5, P7) suggested: '*I would like to have a device that can help me in my first day of class when I meet a lot of unknown classmates, so it can help me remember names and other important aspects such as hometown or age.*'

Device-aesthetics: smartglasses. Participants have some concerns about the aesthetic of the device and prefer a more fashionable wearable, which is also discreet, (Figure 2). The most recalled wearable to use as assistive memory was the *smartglasses*, and also, the most commented solution to visualize the retrieved information was augmented reality (AR). Several participants note an interesting aspect of the wearable device, as P15 (shoes) said: '*I want something I wear and do not need to think of putting on every day.*'

Event-cues: visual and audio. Figure 2 depicts the main interaction methods that participants prefer. Following the device decision, 40% of the participants decide that visual information can help them remember more information about a particular event. Some other interesting participants' opinions (P4, P17) comment: '*I would like a system that can tell me the body language of a person such as gesticulation.*' P5: '*In some situations, I would like a device that can track my biometrical data such as heart-rate so that I can remember my feelings at that moment.*' Several participants (P1, P5, P16) like the idea of audio as a more *private* approach to record events.

Interactions: voice and companion device. Participants decide that voice is a seamless approach to interact with the device, see Figure 2. Participants (P9, P10) highlight that they might be embarrassed by using these voice interactions in public. The main concern with all the interaction methods is the ability to turn the device on-and-off according to their preferences. Participants (P15, P13) choose haptic feedback as a reminder notification technique for future events. Users can interact with VIMES using voice commands, and a companion device such as a smartphone (to access and configure the profile database).

Functionality: summarizing and context-based. Participants (P11, P2) prefer a system that changes its interactions and recording modes according to the user's surroundings, such as conversation topic, location, see Figure 2. P2: '*I would like the system to recognize the context automatically and start collecting the event's information.*' P13: '*I would like the system to recognize the conversation and garner information according to my privacy preferences.*' P8: '*Looking for information in all the collected data can be cumbersome, so I would like to have a system that can provide me the key-points of the events.*' A fire-and-forget usage model will be the default mode to provide accurate and richer information to help users daily.

Additional comments. The user interface was also part of the discussions related to the functionalities and interaction. According to the participants, AR is the best interface approach to show information and display the retrieved data about an event or a person. Our proposed system uses AR and smartglasses to enable users to retain more information and interact with the environment more effectively [22]. Participants are very concerned with the collecting capabilities of the device in private conversations or events. P7: '*Although I would like to have a contact lens device to help in my daily life, I believe the use can be toxic, as it can record all-time.*' However, they would appreciate for the system to be part of a larger life-planning solution. P8: '*I would like the system not only to monitor my surroundings but to get useful information from my online calendar, email, and other online services to provide me with more personalized and accurate information.*' Participants agree that a full-day battery or auto rechargeable one (e.g., user's movement) would be a welcome addition. Other factors were discussed as side-effects of the design choices, such as interacting with records, e.g., P7: '*Timeline and scrolling can be cumbersome when there is a lot of recorded daily information.*' As mentioned earlier, emotions are an essential cue for memory recall [27]. Several participants suggest the idea of recording emotions with VIMES. Modern wearables include electroencephalogram (EEG) sensors that can be used to record users' mental activities such as emotions. *Lifelogging* applications can also collaboratively collect multiple sources of information from different sources, and aggregate them.

Privacy by default. Any system relying on the embedded smartglasses camera captures a wide variety of privacy-sensitive information. Such information may be violated by a third party, depending on where the data is stored and processed. Therefore, privacy preservation must be at the core of the design of the system.

Responsive. The system shall recognize people and return relevant information with a short response delay. Besides, display latency is significantly constrained in mobile AR environments, and should be kept below 20ms for avoiding alignment problems [5].

We focus the prototype of VIMES to audiences with a technology background and some AR experience, as they will be first adopters to use the system. Future developments of the system will focus on audiences that necessary require these memory assistance systems such as dementia patients.

4 SYSTEM ARCHITECTURE

Figure 3 illustrates the system architecture of VIMES, that consists of three modules: visual module, acoustic, and a profile database. We follow: 'the meeting a new person' scenario, as use-case example for the VIMES prototype.

4.1 Recognition Module

The visual module is responsible for detecting faces from images and identifying people using their facial features. In practice, the visual module on the smartglass will detect faces in the camera interface, extract the face features of detected faces, and pass the *feature-vector* (of each face) to the offloading runtime manager. The face processing pipeline is as follows: *(i) face detection* using feature points [49]. *(ii) face recognition*, a query including the received facial feature points (feature-vector), is sent to the companion device's database to retrieve relevant information based on feature matching.

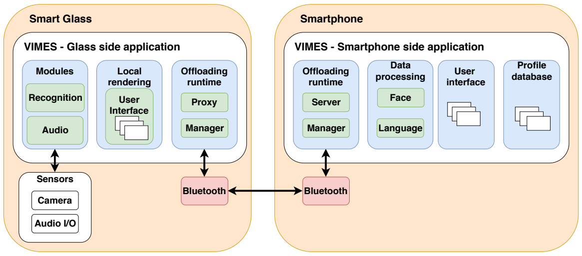


Figure 3: The system architecture of VIMES. Left: Google Glass side. Right: smartphone side.

4.2 Audio Module

The visual module works in conjunction with the audio module, which includes: *(i) speech recognition*, VIMES leverages the Google Glass' speech recognition API [8] to convert speech to text; and *(ii) language processing* extracts useful information from the recognized speech (defined in the profile database). We use OpenNLP for this module, which presents a 90% accuracy [51] to extract real-world information such as persons, locations, and products from unstructured text. The output of speech recognition is passed in text-form to the companion device for further language processing according to the system's profile database configuration.

4.3 Profile Database

The profile database defines how the information will be stored and what information is useful to users daily. The profile database is stored in the companion device as it provides a more feasible interface to manage profiles. The current system uses templates for each type of event that users might find useful to recall. In the 'meeting a new person use case,' the profile database would store facial feature points as a primary key for the *visual module*; personal names, affiliations, and information of previous meetings such as locations and summary of conversations from the *audio module.* The profile data, such as facial features, are stored locally in the smartphone in an encrypted database. We encourage users to use privacy-protective settings during the template creation, such as not enabling by default camera or audio recording (privacy by default). VIMES also emphasizes the minimization of sharing sensitive data such as collected facial features to outside entities, and processes/stores all sensitive information on the companion device in the encrypted profile database (privacy by design).

4.4 Offloading Tasks

We implement a static offloading system between Google Glass and the companion device. The offloading runtime consists of a proxy/server and a manager (see Figure 3). The proxy/server handles the control and data transfer for offloaded tasks. The manager instruments the offloading of face feature points and garnered text from the speech recognition module. We offload the recognition and audio module to the companion device. The devices are paired with Bluetooth as it consumes less energy than WiFi [42] and allow both devices to access the Internet. Our experiments demonstrate that Bluetooth latency is acceptable in our use case.

4.5 System Implementation

We implement VIMES for Google Glass [15] using the Glass Development Kit (GDK). We chose Google Glass as, despite being computationally limited by nowaday's standards, it presents all the sensors and APIs required for our system implementation within a compact form-factor. Besides, in recent years, smartglasses manufacturers have shifted from improving hardware to relying on a companion device for running applications[1]. As such, our prototype system represents a good approximation of more modern systems. We use existing libraries such as OpenCV4Android [38], OpenNLP [39], and SQLite [40]. Our experimental device is Google Glass 2.0 Explorer Edition. The device runs Android 5.0 and is equipped with a 1.2GHz dual-core CPU, a 1GB RAM, and a front camera with its resolution of 1920x1080 pixels. We use a Nexus 4 (Qualcomm APQ8064 Snapdragon S4 Pro Quad-core 1.5 GHz Krait CPU, 2GB RAM) as the companion device. Figure 1 represents the interface on the Google Glass and the companion device.

5 SYSTEM PERFORMANCE

In this section, we evaluate the performance of our system according to response delay, profile retrieval performance, energy consumption, and the accuracy of recognition and audio modules.

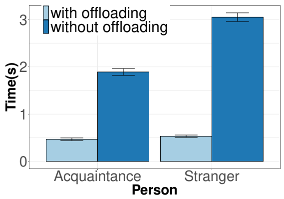
5.1 Experimental Protocol

We evaluate VIMES on the setup described in Section 4.5. All experiments are performed on Google Glass with a companion Nexus 4 smartphone.

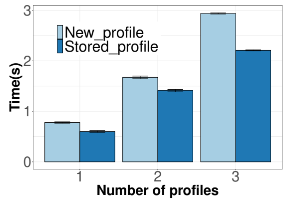
We evaluate the following metrics:

- *Response delay:* the duration between the apparition of a face in the camera frame and the presentation of the corresponding personal information on the display if the profile is already in the system, or the creation of a new profile for a stranger.
- *Profile retrieval performance:* the performance to retrieve profiles information among 100 random generated profiles.
- *Energy consumption:* energy consumed for each device (smartglasses, smartphone) and with/without offloading.
- *Recognition module accuracy:* the face recognition algorithm accuracy in normal sunlight environments. As the distance and view angles could affect the accuracy of this module, a tester using VIMES is placed at various angles and distances from the target. We classify visual module failures into two categories: *(i) Detection failure*; the system does not detect the face in the camera frame. *(ii) Recognition failure*; the system detects the face but does not recognize the person in the profile database.
- *Audio module accuracy:* the percentage of successful recognition (i.e., the module correctly transcribes what the user is saying). The accuracy can vary with different people and different levels of noise. We use prerecorded audio for simulating the user's voice, which varies from 50dB to 70dB[2], while the simulated noise varies from 20dB to 60dB. We generate an audio file with random noise using a random audio noise generator[3].

[1] Vuzix Companion App - https://www.vuzix.com/appstore/app/reflekt-remote
[2] https://www.alpinehearingprotection.co.uk/5-sound-levels-in-decibels/
[3] https://www.random.org/audio-noise/



(a) Response delay for retrieving or creating a single profile.



(b) Profile retrieval performance among 100 profiles.

Figure 4: Average response delay and standard error of the mean) for different scenarios. VIMES benefits greatly from offloading to a companion device (3 to 4 times improvement). VIMES identifies and retrieves profiles in a matter of seconds, which is sufficient for typical human interaction.

We run the experiments 10 times for each metrics and present the average results along with a statistical analysis when applicable.

5.2 Results

Response delay. We evaluate the response delay for a single profile retrieval and creation. Figure 4a illustrates the response delay in both scenarios, with and without offloading. If the user wants to add a stranger to the profile database, both modules (visual and acoustic) must be activated, so the response delay is longer. One-way repeated measures ANOVA ($F(1, 36) = 16.5, p < .001, adjusted - r^2 = 0.95$) shows that the offloading technique has a significant effect on the response delay. The response delay is minimized ($M = 0.45\,s$, 95% : $CI\,[0.39, 0.52]$) when the tasks recognition and language processing are performed on the companion device.

Profile retrieval performance. Figure 4b depicts the profile retrieval performance for multiple faces in the camera frame. The database's new profiles correspond to profiles that are not yet stored in the profile database. One-way repeated measures ANOVA ($F(2, 57) = 276.5, p < .001, adjusted - r^2 = 0.99$) shows a statistically significant effect of the number of profiles stored and the status (new/stored) of the profile. The time to retrieve stored profiles ($M = 1.39\,s$, 95% : $CI\,[1.14, 1.63]$) is significantly shorter than in cases the profile is new and requires to be created ($M = 1.80\,s$, 95% : $CI\,[1.48, 2.12]$). Similarly, the time to retrieve profiles increases when the number of profiles to retrieve increases.

Energy consumption. Besides faster execution times, another benefit of computation offloading comes from the energy consumption reduction of Google Glass. Figure 5 displays over 40% energy savings on Google Glass with computation offloading. The system, including Google Glass and the smartphone, can work together for over 90 minutes. The battery of Google Glass is drained within 45 minutes when running the recognition and audio modules on-device; the 90-min battery life with offloading thus represents a significant improvement and corresponds to the typical battery life of Google Glass with the camera on.

Visual accuracy. Figure 6a depicts the visual module accuracy according to the angle and distance between the user and the individual to recognize. The total visual accuracy declines along with both the distance and angle. The recognition accuracy (face detected but not recognized) is more sensitive to the distance. Within

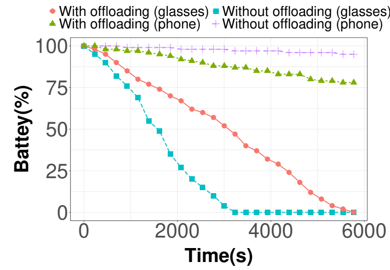
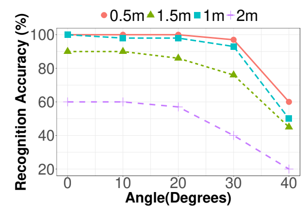
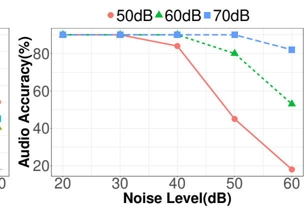


Figure 5: Energy consumption. With offloading, the smartglasses can be used for 90 minutes, as compared to 45 minutes without offloading.



(a) Visual accuracy.



(b) Audio accuracy.

Figure 6: System accuracy. Visual accuracy decreases over 2 m distance and 30°. Audio accuracy is high as long as user voice is 10dB higher than the noise. These values correspond to the requirements for a comfortable conversation.

a two meters range, the visual accuracy of our module is above 90%. The performance then drops significantly (60%). When the individual to retrieve the profile is over 2 meters away from the user, the recognition accuracy significantly decreases as the face resolution falls with the distance. The angle also affects face recognition over 20° to 30°. The system recognizes faces with a high accuracy when they are within the typical field of view of the smartglasses, at conversational distance, which corresponds to our use-case.

Audio accuracy. Figure 6b shows how the audio accuracy drops when the noise level is close to the user's voice. The audio accuracy is 80% when the user's voice level is 60dB, and the noise level is 50dB. In order to keep high audio accuracy, the user should try to maintain the voice level at least 10dB higher than the surrounding noise, which is socially expected in a normal conversation.

6 USER EVALUATION

After evaluating the system performance, we proceed to measure the users' performance and their subjective perception of VIMES.

6.1 Participants and Apparatus

To understand users' memory recall performance, memory experience, awareness, and intended adoption, we recruit 50 participants. 53% are male, 47% female, with ages ranging from 19 to 31 (M = 24, SD = 3.07). The participants are students invited from business (20%), science (30%), and engineering (50%) schools at our University. All participants noted that they have little prior experience of using smartglasses like Google Glass. There was no overlap between the evaluation group and the co-design study group.

The experiment was carried out in accordance with the local IRB regulations. We informed the participants that the data will be password-protected and deleted after the study ends, and participants provided informed consent. They were rewarded with sweets and soft drinks after completion of the experiment.

6.2 Experimental Protocol

We evaluate the memory retrieval performance of our proposed system against two traditional memory recall methods as follows:

Self-recall. Individuals retrieve information without aid. Users retrieve the information only using their own memory capabilities.

Visual-cues. Individuals are helped with visual cues (i.e., faces of persons) to help them retrieve the asked information.

VIMES. Individuals will use our system to retrieve the information.

In this user study, we focus on the following metrics

- *Memory performance.* We conduct a memory test to quantitatively evaluate (recall performance score, RPS) how the system can provide memory recall to users. We evaluate the described recall methods. We ask participants to recall the time, location, name, background, and opinion of the other three students measure their binary recall performance (0: fail, 1: success).
- *Memory experience.* We ask participants their *memory experience* [31]: vividness, coherence, accessibility, and visual perspective using a 5-point scale (Answer: 1.lowest - 5.highest).
- *Technology Acceptance.* We follow the technology acceptance model (TAM), which includes two criteria: perceived usefulness (PU), and perceived ease of use (PEOU) We follow a 5-point scale to quantitatively measure the two criteria (Answer: 1.strongly disagree - 5.strongly agree).
- *Individual awareness and intended adoption.* We first evaluate participants' privacy concerns using the Internet User's Information Privacy Concerns (IUIPC) [32]. The IUIPC focus on the collection, awareness, and control of data. We then analyze the different perceptions of awareness and intended adoption from both the user and the bystander. Participants answer questions using a 5-point scale (Answer: 1.very unlikely - 5.very likely).

To evaluate participants' memory performance, we design a two-stage experiment. First, users meet and discuss with three unknown students. In a second meeting, we ask them to recall information about their encounter. Time hinders memory recall [58, 59]. To provide a conservative comparison, we run this test only one day after the first meeting.

First meeting. Each participant has a meeting with the same three unknown students. Together, the participant and the students introduce themselves and discuss the following topic: *Can alternative energy effectively replace fossil fuels?* for ten minutes. We choose this topic as to keep participants engaged in the discussion. The participant wears the VIMES device during the meeting. With this discussion approach we aim to evaluate the personal information retrieval performance and how much information participants can remember from other students' point of views about the topic.

One day later, we ask each participant to recall information about the three students, including the time and location, the students' names and background, and their opinions. We distribute the counterbalanced questions for each memory recall method. Therefore, we do not ask the same question twice using two different memory recall methods. We counterbalance the memory recall methods . We conclude with a survey to evaluate the users' subjective perception of the system, including their *memory experience*, the TAM questionnaire, and their awareness and intended adoption.

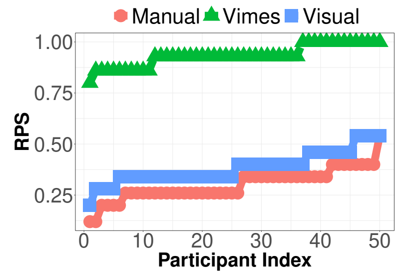


Figure 7: Distribution of recall performance score (RPS). VIMES improves the RPS by 30 to 100%.

Table 1: Users' Rating of Memory Experience.

Memory Experience (# items)	**Self Recall** mean (SD)	**Visual Cues** mean(SD)	**VIMES** mean(SD)
Vividness (3)	3.1 (0.74)	3.4 (0.58)	4 (0.67)
Coherence (4)	2.7 (0.48)	2.9 (0.52)	4.1 (0.57)
Accessibility (3)	2.9 (0.74)	3.1 (0.76)	3.8 (0.63)
Visual Perspective (3)	3 (0.67)	3.6 (0.469)	4.1 (0.57)

6.3 Results

Memory Recall Performance (percentage). Figure 7 depicts the distribution of RPS of all participants. One-way repeated measures ANOVA $F(2, 98) = 74.78, p < .001, adjusted - r^2 = 0.9416$) shows that the memory retrieval method has a significant effect on memory recall performance. The use of VIMES for retrieving information shows an improvement ($M = 0.93$, $95\% : CI\,[0.92, 0.95]$), compared to self-retrieval ($M = 0.30$, $95\% : CI\,[0.28, 0.32]$) or visual-cues ($M = 0.39$, $95\% : CI\,[0.36, 0.4]$). As other works stated [28, 59], wearable cameras improve users' autobiographical memory recall.

Memory Experience. Table 1 shows participants' average ranking on the factors of memory experience. When using the visual method or VIMES, the average score of the system is four, while the average manual review score is three. For VIMES, one-way repeated measures ANOVA indicates a statistically significant difference for: Vividness, $F(2, 98) = 25.11, p < .001, adjusted - r^2 = 0.26$; Coherence, $F(2, 98) = 18.85, p < .001, adjusted - r^2 = 0.19$; Accessibility, $F(2, 98) = 34.84, p < .001, adjusted - r^2 = 0.31$; Visual perspective, $F(2, 98) = 41.75, p < .001, adjusted - r^2 = 0.35$. All the participants report that they can better recall the meeting's events with the assistance of the system.

User Survey. In our TAM survey, participants rate high the perceived usefulness of the device ($M = 3.75$, $95\% : CI\,[2.83, 4.66]$) and the perceived ease of use ($M = 3.46$, $95\% : CI\,[2.60, 4.33]$). Participants also left many comments during the experiment. Most noted that operating VIMES is a unique, refreshing, and user-friendly experience. Many participants show interest in seeing such an application in the market in the future. Example of specific comments include *P23*: '*I like the idea, and it would be beneficial to me because I often forget the names of the students in my class.*' P27: '*I usually take photos or record voices as a reminder but have never tried to combine these two ways. This idea provides a convenient way which could be quite useful.*' P42: '*Interesting experiment but sometimes the Glass can be intrusive. I would worry about my privacy if I need to wear it for a long time.*' Overall, most users are satisfied with the application, and they highlight the convenience and potential usefulness of the system.

Individual awareness and intended adoption. Participants are very concerned about privacy in their online behavior. However, most of them admit they use services if the benefits are higher than the privacy threats. To analyze the IUIPC, we first perform Principal Component Analysis (PCA) to verify each scale's dimensionality. The IUIPC PCA shows the four original components predicted the total variance: collection ($\alpha = 0.877$), control ($\alpha = 0.76$), awareness ($\alpha = 0.8$), and error ($\alpha = 0.68$). 80% of participants prefer a discreet device as users, but, as bystanders, the opinions are distributed along the 5-point scale. Most participants are very comfortable (33%) or neither comfortable nor uncomfortable (40%) asking VIMES users to delete the recorded information. More than 90% of participants believe that VIMES users should ask for permission to record.

7 DISCUSSION AND FUTURE WORK

Technical Performance. VIMES presents a high performance in the scope of face-to-face and small-group conversations. It detects and recognizes faces with high accuracy, up to 2 meters, and 30°. The audio accuracy is over 80% for voices 10dB higher than the noise. For up to three individuals in the user's field of view, VIMES extracts or creates profiles in less than three seconds, which is acceptable in day-to-day interactions.

User Acceptance and Intended Adoption. Most participants agreed that the wearable device should be discrete, fashionable, and essential. Smartglasses fit these requirements. VIMES' summarizing capabilities are very welcome to deal with the myriad of information garnered and highlight the key aspects of an event. VIMES improves users' autobiographical memory with high usability. One of the main adoption barriers of VIMES is the privacy issues that arise when someone else is using it.

Hardware Limitations. Our prototype relies on *Google Glass*, a 2013 piece of hardware. However, most recent smartglasses only perform sensing and display while a companion smartphone runs the applications. As such, our prototype system is representative of more recent smartglasses' operation. Besides, it ensures that our algorithms run on every available piece of hardware. With only 90 minutes, *Battery life* is another issue. The camera is one of the most energy-hungry components of smartglasses. Recent smartglasses present more energy-efficient cameras and larger batteries. Context detection (e.g., audio [41], video [62]) to activate sensors automatically can also significantly improve the battery life.

Software Limitations. *Interacting with larger groups of people* is a challenge in the current implementation. The association of voice and face in multiple-persons environments is a challenging problem [13]. Besides, the recognition of multiple faces in images is still a challenging topic that limits the speed to recognize multiple faces [18]. Larger groups also increase the physical distance between participants, over 2m/30°. Recent smartglasses feature cameras with higher resolution and field of view, thus circumventing the problem. Our prototype system relies on a *single hard-coded topic. Topic modeling* [11] would allow us to extract information from the conversation according to the context. Our current summarization of egocentric video using templates distant from more flexible summarization techniques in-the-wild such as [43].

Privacy is a common concern among participants. Our privacy by design approach requires VIMES to keep all sensitive information on-device. Homomorphic encryption would present an additional security layer to prevent an eventual attacker from obtaining information from a compromised device. The device-to-device communication protocol should also follow security and privacy standards to protect the transmitted information between the smart glasses and companion devices. For privacy awareness and preservation, we should follow novel notification methods to inform bystanders about the monitoring of VIMES [30], input interfaces such as gestures [29, 54] to opt-in/opt-out consent, and privacy protection techniques when bystanders opt-out [53].

Social Acceptance. Besides privacy, smartglasses and AR are subject to social acceptance issues. Interaction methods are often disruptive, whether gestural or vocal [56]. By relying heavily on face recognition and the companion device in VIMES, we limit the user interaction cost considerably. Non-users may also feel uncomfortable or offended by users relying on an external system to remember facts about them. However, with the democratization of smartglasses and AR, we expect such concerns to fade away.

8 CONCLUSION

In this work, we designed, implemented, and evaluated a wearable AR system to help recall a memory of facts and events. Participants in our collaborative design study think that smartglasses are the best peripheral for autobiographical memory aid. The proposed system follows participants' functionality requests such as a good battery (task offloading doubles the battery life), summarization techniques, and a companion device to provide further interactions. We perform several experiments to evaluate the system both in terms of performance and user perception. VIMES generates and retrieves personal profiles with an average accuracy of over 90% in practical scenarios, and users find VIMES not only useful but also easy to use. The user study also highlights the future challenges the adoption of such a system will encounter. With the development of smartglasses hardware and iterative improvements in our system, we believe it offers excellent potential for improving memory recall in daily activities.

9 ACKNOWLEDGEMENTS-PLACEHOLDER

This research has been supported in part by project 16214817 from the Research Grants Council of Hong Kong,and the 5GEAR and FIT projects from Academy of Finland.

REFERENCES

[1] Mousa Ahmadi, Cristian Borcea, and Quentin Jones. 2019. Collaborative lifelogging through the integration of machine and human computation. In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*. 23–24.

[2] Mélissa C Allé, Liliann Manning, Jevita Potheegadoo, Romain Coutelle, Jean-Marie Danion, and Fabrice Berna. 2017. Wearable cameras are useful tools to investigate and remediate autobiographical memory impairment: A systematic PRISMA review. *Neuropsychology review* 27, 1 (2017), 81–99.

[3] Marc Bächlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M Hausdorff, Nir Giladi, and Gerhard Tröster. 2010. Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. *Information Technology in Biomedicine, IEEE Transactions on* 14, 2 (2010), 436–446.

[4] Michelle S Bourgeois. 1993. Effects of memory aids on the dyadic conversations of individuals with dementia. *Journal of Applied Behavior Analysis* 26, 1 (1993), 77–87.

[5] Tristan Braud, Farshid Hassani Bijarbooneh, Dimitris Chatzopoulos, and Pan Hui. 2017. Future networking challenges: The case of mobile augmented reality. In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 1796–1807.

[6] Georgina Browne, Emma Berry, Narinder Kapur, Steve Hodges, Gavin Smyth, Peter Watson, and Ken Wood. 2011. SenseCam improves memory for recent events and quality of life in a patient with memory retrieval difficulties. *Memory* 19, 7 (2011), 713–722.

[7] Laurindo de Sousa Britto Neto, Vanessa Regina Margareth Lima Maike, Fernando Luiz Koch, Maria Cecília Calani Baranauskas, Anderson de Rezende Rocha, and Siome Klein Goldenstein. 2015. *A Wearable Face Recognition System Built into a Smartwatch and the Blind and Low Vision Users*. Springer International Publishing, Cham, 515–528. https://doi.org/10.1007/978-3-319-29133-8_25

[8] Google Developers. 2015. https://developers.google.com/glass/develop/ gdk/voice.

[9] Lina Dib, Daniela Petrelli, and Steve Whittaker. 2010. Sonic souvenirs: exploring the paradoxes of recorded sound for family remembering. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, 391–400.

[10] Aiden R Doherty and Alan F Smeaton. 2008. Combining face detection and novelty to identify important events in a visual lifelog. In *Computer and Information Technology Workshops, 2008. CIT Workshops 2008. IEEE 8th International Conference on*. IEEE, 348–353.

[11] Mortaza Doulaty, Oscar Saz, and Thomas Hain. 2015. Unsupervised domain discovery using latent dirichlet allocation for acoustic modelling in speech recognition. *arXiv preprint arXiv:1509.02412* (2015).

[12] eMarketer. 2017. Wearable user penetration rate in the United States, in 2017, by age. https://www.statista.com/statistics/739398/us-wearable-penetration-by-age/. [Online; accessed 4-February-2019].

[13] Ariel Ephrat, Inbar Mosseri, Oran Lang, Tali Dekel, Kevin Wilson, Avinatan Hassidim, William T Freeman, and Michael Rubinstein. 2018. Looking to listen at the cocktail party: A speaker-independent audio-visual model for speech separation. *arXiv preprint arXiv:1804.03619* (2018).

[14] Mohammad Ghafouri, Shohreh Amini, Kamel Khalili, and Bassel E Sawaya. 2006. HIV-1 associated dementia: symptoms and causes. *Retrovirology* 3, 1 (2006), 28.

[15] Google Glass. 2015. https://developers.google.com/glass.

[16] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Foundations and Trends in Information Retrieval* 8, 1 (2014), 1–125. https://doi.org/10.1561/1500000033

[17] Kiryong Ha, Zhuo Chen, Wenlu Hu, Wolfgang Richter, Padmanabhan Pillai, and Mahadev Satyanarayanan. 2014. Towards wearable cognitive assistance. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. ACM, 68–81.

[18] Catrina M Hacker, Emily X Meschke, and Irving Biederman. 2019. A face in a (temporal) crowd. *Vision research* 157 (2019), 55–60.

[19] Elmar Hashimov. 2015. Qualitative Data Analysis: A Methods Sourcebook and The Coding Manual for Qualitative Researchers: Matthew B. Miles, A. Michael Huberman, and Johnny Saldaña. Thousand Oaks, CA: SAGE, 2014. 381 pp. Johnny Saldaña. Thousand Oaks, CA: SAGE, 2013. 303 pp.

[20] Steve Hodges, Emma Berry, and Ken Wood. 2011. SenseCam: A wearable camera that stimulates and rehabilitates autobiographical memory. *Memory* 19, 7 (2011), 685–696.

[21] Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Ken Wood. 2006. SenseCam: A retrospective memory aid. In *UbiComp 2006: Ubiquitous Computing*. Springer, 177–193.

[22] Lei Hou and Xiangyu Wang. 2013. A study on the benefits of augmented reality in retaining working memory in assembly tasks: A focus on differences in gender. *Automation in Construction* 32 (2013), 38–45.

[23] Zhanpeng Huang, Weikai Li, and Pan Hui. 2015. Ubii: Towards Seamless Interaction between Digital and Physical Worlds. In *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference*. ACM, 341–350.

[24] Sinh Huynh, Rajesh Krishna Balan, and Youngki Lee. 2015. Demo: Towards Recognition of Rich Non-Negative Emotions Using Daily Wearable Devices. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*. ACM, 471–472.

[25] Ellen Isaacs, Artie Konrad, Alan Walendowski, Thomas Lennig, Victoria Hollis, and Steve Whittaker. 2013. Echoes from the past: how technology mediated reflection improves well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1071–1080.

[26] Katherine Isbister, Kaho Abe, and Michael Karlesky. 2017. Interdependent Wearables (for Play): A Strong Concept for Design.. In *CHI*. 465–471.

[27] Shiqi Jiang, Zhenjiang Li, Pengfei Zhou, and Mo Li. 2019. Memento: An emotion-driven lifelogging system with wearables. *ACM Transactions on Sensor Networks (TOSN)* 15, 1 (2019), 1–23.

[28] Vaiva Kalnikaite, Abigail Sellen, Steve Whittaker, and David Kirk. 2010. Now let me see where i was: understanding how lifelogs mediate memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2045–2054.

[29] Marion Koelle, Swamy Ananthanarayan, Simon Czupalla, Wilko Heuten, and Susanne Boll. 2018. Your smart glasses' camera bothers me! exploring opt-in and opt-out gestures for privacy mediation. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. 473–481.

[30] Marion Koelle, Katrin Wolf, and Susanne Boll. 2018. Beyond LED status lights-design requirements of privacy notices for body-worn cameras. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. 177–187.

[31] Martina Luchetti and Angelina R Sutin. 2016. Measuring the phenomenology of autobiographical memory: A short form of the Memory Experiences Questionnaire. *Memory* 24, 5 (2016), 592–602.

[32] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[33] Steve Mann. 1997. Wearable computing: A first step toward personal imaging. *Computer* 30, 2 (1997), 25–32.

[34] Daniel McDuff, Amy Karlson, Ashish Kapoor, Asta Roseway, and Mary Czerwinski. 2012. AffectAura: an intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 849–858.

[35] A McPherson, FG Furniss, C Sdogati, F Cesaroni, B Tartaglini, and J Lindesay. 2001. Effects of individualized memory aids on the conversation of persons with severe dementia: a pilot study. *Aging & Mental Health* 5, 3 (2001), 289–294.

[36] Pranav Misra, Alyssa Marconi, Matthew Peterson, and Gabriel Kreiman. 2018. Minimal memory for details in real life events. *Scientific reports* 8, 1 (2018), 1–11.

[37] Vivian Genaro Motti and Kelly Caine. 2015. Users' privacy concerns about wearables. In *International Conference on Financial Cryptography and Data Security*. Springer, 231–244.

[38] OpenCV4Android. 2016. http://opencv.org/platforms/android.html.

[39] Apache OpenNLP. 2010. https://opennlp.apache.org.

[40] Mike Owens and Grant Allen. 2010. *The Definitive Guide to SQLite*. Springer.

[41] Giambattista Parascandolo, Heikki Huttunen, and Tuomas Virtanen. 2016. Recurrent neural networks for polyphonic sound event detection in real life recordings. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6440–6444.

[42] Trevor Pering, Yuvraj Agarwal, Rajesh Gupta, and Roy Want. 2006. Coolspots: reducing the power consumption of wireless mobile devices with multiple radio interfaces. In *Proceedings of the 4th international conference on Mobile systems, applications and services*. ACM, 220–232.

[43] Anuj Rathore, Pravin Nagar, Chetan Arora, and CV Jawahar. 2019. Generating 1 Minute Summaries of Day Long Egocentric Videos. In *Proceedings of the 27th ACM International Conference on Multimedia*. 2305–2313.

[44] Philipp A Rauschnabel, Alexander Brem, and Bjoern S Ivens. 2015. Who will buy smart glasses? Empirical results of two pre-market-entry studies on the role of personality in individual awareness and intended adoption of Google Glass wearables. *Computers in Human Behavior* 49 (2015), 635–647.

[45] Philipp A Rauschnabel and Young K Ro. 2016. Augmented reality smart glasses: An investigation of technology acceptance drivers. *International Journal of Technology Marketing* 11, 2 (2016), 123–148.

[46] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.

[47] Corina Sas, Scott Challioner, Christopher Clarke, Ross Wilson, Alina Coman, Sarah Clinch, Mike Harding, and Nigel Davies. 2015. Self-defining memory cues: creative expression and emotional meaning. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013–2018.

[48] Corina Sas, Tomasz Fratczak, Matthew Rees, Hans Gellersen, Vaiva Kalnikaite, Alina Coman, and Kristina Höök. 2013. AffectCam: arousal-augmented sensecam for richer recall of episodic memories. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1041–1046.

[49] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE*

conference on computer vision and pattern recognition. 815–823.

[50] Abigail J Sellen, Andrew Fogg, Mike Aitken, Steve Hodges, Carsten Rother, and Ken Wood. 2007. Do life-logging technologies support memory for the past?: an experimental study using sensecam. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 81–90.

[51] Khaled Shaalan. 2014. A survey of arabic named entity recognition and classification. *Computational Linguistics* 40, 2 (2014), 469–510.

[52] Yiran Shen, Chengwen Luo, Weitao Xu, and Wen Hu. 2015. Poster: An Online Approach for Gait Recognition on Smart Glasses. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*. ACM, 389–390.

[53] Zhiqi Shen, Shaojing Fan, Yongkang Wong, Tian-Tsong Ng, and Mohan Kankanhalli. 2019. Human-imperceptible Privacy Protection Against Machines. In *Proceedings of the 27th ACM International Conference on Multimedia*. 1119–1128.

[54] Jiayu Shu, Rui Zheng, and Pan Hui. 2016. Cardea: Context-Aware Visual Privacy Protection from Pervasive Cameras. *eprint arXiv:1610.00889* (Oct 2016), 1–10.

[55] Enrico Tanuwidjaja, Derek Huynh, Kirsten Koa, Calvin Nguyen, Churen Shao, Patrick Torbett, Colleen Emmenegger, and Nadir Weibel. 2014. Chroma: A wearable augmented-reality solution for color blindness. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 799–810.

[56] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike Y Chen. 2015. User-defined game input for smart glasses in public space. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3327–3336.

[57] Dapper Vision. 2014. http://www.openshades.com.

[58] Willem A Wagenaar. 1986. My memory: A study of autobiographical memory over six years. *Cognitive psychology* 18, 2 (1986), 225–252.

[59] Emma Woodberry, Georgina Browne, Steve Hodges, Peter Watson, Narinder Kapur, and Ken Woodberry. 2015. The use of a wearable camera improves autobiographical memory in patients with Alzheimer's disease. *Memory* 23, 3 (2015), 340–349.

[60] WHO World Health Organization. 2017. Dementia. http://www.who.int/news-room/fact-sheets/detail/dementia. [Online; accessed 15-October-2018].

[61] Y. Wu, X. Shen, T. Mei, X. Tian, N. Yu, and Y. Rui. 2016. Monet: A System for Reliving Your Memories by Theme-Based Photo Storytelling. *IEEE Transactions on Multimedia* 18, 11 (Nov 2016), 2206–2216. https://doi.org/10.1109/TMM.2016.2614185

[62] Yingying Zhu, Nandita M Nayak, and Amit K Roy-Chowdhury. 2013. Context-aware modeling and recognition of activities in video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2491–2498.