Accepted Manuscript

Remembering through lifelogging: A survey of human memory augmentation

Morgan Harvey, Marc Langheinrich, Geoff Ward

PII: S1574-1192(15)00214-X
DOI: http://dx.doi.org/10.1016/j.pmcj.2015.12.002
Reference: PMCJ 664

To appear in: *Pervasive and Mobile Computing*

Received date: 19 March 2015
Revised date: 24 November 2015
Accepted date: 9 December 2015

Remembering through lifelogging: a survey of human memory augmentation

Morgan Harvey[a,*], Marc Langheinrich[b], Geoff Ward[c]

[a]Department of Maths and Information Sciences, Northumbria University, Newcastle, UK
morgan.harvey@northumbria.ac.uk, Tel: +44 (0)191 349 5929
[b]Faculty of Informatics, Università della Svizzera italiana (USI), Switzerland
[c]Department of Psychology, University of Essex, UK

Abstract

Human memory is unquestionably a vital cognitive ability but one that can often be unreliable. External memory aids such as diaries, photos, alarms and calendars are often employed to assist in remembering important events in our past and future. The recent trend for lifelogging, continuously documenting ones life through wearable sensors and cameras, presents a clear opportunity to augment human memory beyond simple reminders and actually improve its capacity to remember. This article surveys work from the fields of computer science and psychology to understand the potential for such augmentation, the technologies necessary for realising this opportunity and to investigate what the possible benefits and ethical pitfalls of using such technology might be.

Keywords: Lifelogging, Augmented Human Memory, Personal Life Archives

1. Introduction

Human memory is a critical cognitive function that we rely on almost constantly in our everyday lives. External memory aids are often used to help support memory for past events; photographs help us to remember autobiographical events such as holidays, we use recorded minutes to help us remember the content of meetings, and revision notes to remember lectures. Similarly, post-it notes, alarms, and interactive calendars and reminders help us to remember future events. Such aids often require some planning and conscious effort to initiate and record, and provide external support for only a small proportion of all past and future events. Thus we often must rely on our human memory to encode, store, and later retrieve our experiences. Although our minds are often able to recollect information quickly upon request, we know that our memory may fail us, particularly as we age.

*Corresponding author

In this article, we consider the potential of using lifelogging technology to mitigate against everyday memory failure. In keeping with the descriptions of lifelogging offered by Dodge and Kitchin [1], and then more recently by Gurrin et al. [2], we consider lifelogging to be a "form of pervasive computing, consisting of a unified digital record of the totality of an individuals experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive". We envisage that the recorded lifelog data of a person's life could not only be used as an enhanced everyday memory aid or memory prosthesis [3], but through the re-presentation and review of such data, could also be used to enhance or augment human memory. We first review the psychology of human memory literature to assess the potential for the use of such technologies and summarise early research on memory problems and memory aids. We discuss contemporary lifelogging technologies, identify key sources of data that can be recorded and describe ways that this raw data can first be segmented into manageable events or "episodes" and then further processed to turn it into useful information. We investigate how the various forms of recorded information (e.g., images, videos and text) can be effectively organised, managed and searched in such a manner that it builds upon the natural tendencies of the brain. We consider the privacy and ethical issues that the everyday use of this technology might raise. Finally we discuss ways in which all of these technologies could be brought together in the near future to support and augment memory.

2. Human memory and the potential for augmentation

Although the existence of strict laws or principles of memory is subject to debate [4], we can still outline three key features of human memory that, when considered together, demonstrate the potential for human memory augmentation.

First, research has demonstrated that our ability to accurately recall any given item or episode from human memory is likely to be relatively poor because any particular item is likely to be similar to earlier or later items. As such the task of memory retrieval requires the discrimination of the target item from its competitors: recall will depend upon the relative distinctiveness of the target event when compared to its competitors [4]. As the number of subsequent similar events increases, so the ability to retrieve a target event decreases (retroactive interference). Likewise, as the number of similar prior events increases, so too does the difficulty in retrieving the target event (proactive interference) [5]. The relative importance of retroactive and proactive interference depends upon the temporal separation between the target and interfering memories and also the retention interval (the time between study of the target and the memory test), as predicted by temporal or contextual discrimination theory [6]. These phenomena are observed in word lists and also in more "real-world" eye-witness testimony methodologies.

Second, human memory is cue driven [4] and these cues can be verbal, images, nonverbal sounds, emotions or mood, locations and study environment [7]. A cue is effective if it is specific and if it was present during the encoding of the item [8]. Therefore a re-presented item that has both of these characteristics is likely to be relatively easy to recognise, and in most circumstances we can recognise stimuli far better than we can recall them [9]. Lifelogging has the potential to capture data that may be useful as a retrieval cue: a holiday image may not only be easily recognised, but it might serve as a specific cue to help retrieve other associated or temporally contiguous memories related to that event. It is equally clear that effective cues could be derived from multiple media types (or their combination).

Third, cues captured by lifelogging technology can be used to augment human memory. Events that we wish to remember better can be re-presented to us for later review and testing. Not only can the spontaneous recall of items be improved through spaced presentations, but trying to remember something can actually help cement things in memory more effectively than further study (the "testing effect") [10]. Moreover, while retrieval practice of events can improve their subsequent retention, there is evidence that non-practised items related to those that were practised can be actively inhibited (retrieval-induced forgetting) [11]. The potential for lifelogging technology to augment spontaneous recall by providing cues to memories for subsequent revision sessions is immediately apparent, and there is the intriguing possibility that such technology could also actively suppress the recall of unwanted memories.

2.1. Memory problems

Several early diary studies [3, 12] have shown that the memory problems that people encounter in their everyday lives are both prospective (e.g. forgetting to meet someone or do something) and retrospective in nature (forgetting a name or phone number, a locations, or facts about other people) [13]. Elsweiler et al. [14] investigated typical everyday memory problems by conducting a diary study over the period of a week with 25 participants who were asked to keep a log of any memory lapses they experienced. Although the majority (51%) of the 261 recorded lapses were retrospective, a large minority (38%) were prospective.

A more recent one week diary study by Unsworth et al. [15] of 100 students that focused on failures of retrospective events showed that forgetting information for an exam or homework was the highest reported failure, with (in descending order of frequency) failure to recall names, password or login ID, friend or family information, directions, and non-college facts the next most frequent. Interestingly, even in this study there were failures of a prospective nature (forgetting what one was doing or looking for and forgetting what one was going to tell someone), a feature that is commonly observed in more open-ended diary studies.

Research has shown [16] that a lack of attention paid during an event results in the probability of said event being recalled accurately, or indeed at all, being significantly reduced. Likewise, the amount of processing performed on the stimulus at the time it is experienced has a large impact on its future retrievability [17]. Indeed, many instances of forgetting can be put down to the ineffective *encoding* - how the new memory is integrated into the pool of existing memories - of the experience while it is taking place, often as a result of one's attention being divided and thus not being able to focus clearly on the event itself.

The potential for lifelogging technology to act as a surrogate memory or prosthesis is therefore immense. It could allow the user to circumvent many of these everyday memory failures. Since it is automatic and cannot be "distracted," lifelogging can help overcome failures in encoding, leaving the user to concentrate on the task of remembering clues to cue the memory. If sufficient information about each memory is made available (and a user interface is provided allowing the user to efficiently explore the memory space), then the amount of contextual information at hand can be greater. The user could restudy data captured during earlier events and encode for the first time aspects of the event that were missed, or be reminded through the captured data of aspects of the events that had temporarily been inaccessible through changes in context or goals. Digital images are likely to provide particularly potent cues to elicit memories of previous experiences. It is already well established that reviewing static images from a previously viewed movie of everyday events can help subsequent free and cued recall and recognition [18].

Perhaps the most important single piece of lifelogging technology used in the psychology of memory research to date is SenseCam [19], a small device fitted with sensors and a digital camera worn as a pendant around the neck. Developed by Microsoft, it takes still images automatically while it is worn, and features a fish-eye lens allowing most of what the wearer sees to be captured, including people they speak to, and removing the burden on the user of framing a shot. By default, the device take images every 30 seconds, but detected changes in light levels and movement can also be used to trigger image capture. Stored images are time stamped and location data can be recorded by a GPS receiver (although this feature is not standard). Other features include the manual activation of the camera by the user, as in a normally functioning camera, and a privacy button that suspends image capture for four minutes.

Custom software allows the downloaded images to be viewed individually or in a speeded sequence analogous to a movie and research has shown that reviewing the images taken can significantly improve subsequent recall performance in both healthy [20] and memory-impaired users [21, 22]. Hodges et al. [23] believe that the SenseCam images provide effective retrieval cues to personal experiences, and can often elicit thoughts, feelings and emotions from the time of the event. It is likely that there will be at least one cue to what one is trying to remember from one's day is within the temporally-ordered sequences images that taken that day. SenseCam's effectiveness in augmenting memory has typically relied upon compiling images from the salient events of the day into an end of day review.

It is this summary review that has perhaps the greatest unexplored potential to be exploited for augmenting human memory. These previous studies have assumed that reviewing a subset of material at the end of the day will improve subsequent retention, especially if the reviewed material is used as a retrieval cue to prompt actual memories of the event. These assumptions are completely compatible with the well known increase in recall attributed to increased repetition, recency, and distribution of rehearsals encountered in the laboratory [24]. The assumptions are also compatible with the memorial advantage observed with testing memory on later retention (e.g., testing effects [10]) and the benefits of retrieval practice [11].

However, none of the lifelogging studies to date have considered another recent theoretical development from laboratory studies within the psychology of memory, that the selective recall of a subset of events can actual impair accessibility to related but unreviewed events (a phenomenon known as retrieval induced forgetting, RIF [11]). To date, the vast majority of RIF studies have been conducted within the laboratory, very few have been conducted under more real world conditions. Moreover, none of the lifelogging studies that have employed an end-of-day review, nor laboratory studies that have considered RIF have systematically examined the factors will enhance the magnitude of retrieval practice of the reviewed items and the magnitude of retrieval induced forgetting of the related but unreviewed items. An understanding of these factors is important as participants are likely to wish the maximum memorial benefit from a minimum period of review.

3. Logging a life

Although the concept of lifelogging has in principle been around for over 30 years [25], practical examples have had to wait for technology to develop. Some of the earliest examples of lifelogging platforms were developed by Steve Mann [26, 27] and were originally somewhat unwieldy, requiring the user to wear a large helmet and a belt-mounted battery pack. While this may have been suitable for research purposes, people are likely to be reluctant to use such a device on an everyday basis. Later revisions of Mann's work involved integrating the lifelogging device into apparel such as glasses [28], an approach which has become more common recently due to the introduction of more mainstream devices such as Google Glass or the Epson Moverio. Perhaps the most frequently used example of a lifelogging device is Microsoft's SenseCam [19], created in 2006. SenseCam technology has since been licensed to a company and re-appeared in improved versions first as the "Vicon Revue" and more recently as the "OMG Autographer" (discussed later).

While lifelogging often focuses on the recording of images or even video, the general concept applies to all kinds of personal data sources, such as one's position as recorded via GPS, audio recordings, and of course physiological data such as heart rate or step counters [29, 19]. Our "digital traces" can also be sources for lifelogging, such as emails, calendar events, documents, instant messages, and items posted on social media. The key similarity of these sources is that they continuously and passively (i.e. automatically) record a user's personal experiences. In this article we predominantly focus on image capture devices. Episodic memory is innately dominated by visual imaginary, and visual cues (e.g. pictures or videos) are known for maximising the information they contain by representing objects in relation to each other [30]. For a detailed survey of the history of lifelogging the reader is referred to work by Gurrin et al. [2].

A lifelogging device design to capture the visual context of one's life must be designed with certain constraints in mind. It should not be too bulky or heavy, as it will generally be worn on the head (e.g., as glasses) or around the neck (e.g., on a lanyard) for long periods of time. Ideally the wearer should forget about its presence so that they behave naturally and forget that their actions are being recorded. Being small, the device can also be used in a less conspicuous manner as other people are not distracted by the user wearing the device.[1] Battery life is a salient factor as having to frequently stop using the device to allow it to charge would not only be tiresome but would limit the amount of data that could be captured. Despite the practical limitations of current battery technology, the amount of data collected by modern image capture systems is still vast and can yield useful information, provided the raw images collected are intelligently processed [31]. Today's visual lifelogging devices not only record images, but often also other forms of contextual information, such as the time of capture, the GPS position of the wearer, light levels, and even acceleration (to orient the image). All of these contextual features can potentially be used in concert with the images to increase the likelihood that memory will be successfully jogged by the information.

Depending on the intended use of the data recorded, perhaps the most important characteristic is the quality of images captured, especially if the resolution should be sufficient to be able to make out small details such as people's faces. It is also crucial that the camera system has a fast shutter speed (meaning a wide aperture or sensitive light sensor), since it will often be in motion when images are taken. Some lifelogging systems feature a fish-eye lens, as this allows more of the wearer's surroundings to be be captured in each image. However, this in turn carries the drawback that objects in images can become heavily distorted.

A number products have become available in recent years that feature many if

[1]This clashes with the principle of avoiding covert recording of others. See section 3.1.

Figure 1: Three lifelogging devices: the Autographer, Narrative Clip and SenseCam.

not all of these desirable properties. The most prominent devices are probably the "Autographer" and the "Narrative Clip" [32] (left and centre images of figure 1). The Autographer was developed based on technology licensed from the SenseCam project and represents a logical continuation of it. It is smaller and lighter, features a 5 megapixel camera with a 136-degree wide-angle lens and can take up to 2,000 photos a day which are stored in its 8GB internal memory. It also has sensors to detect changing light levels, temperature, motion and orientation, as well as a GPS receiver ensuring that photos are geo-tagged. The Narrative Clip instead was developed independently with a slightly different focus in mind. In contrast to the distorted images coming from a fisheye lens, its moderate wide-angle lens (70-degree) provides "normal" pictures that are comparable to a mobile phone camera. While this means that it is unable to capture quite as much of the wearer's surroundings, its pictures are much more "shareable" with others. It is also significantly smaller than the Autographer, making it not only easier to wear but also much more unobtrusive. However, as a result of its miniaturisation, the Clip not only has fewer sensors, but also lacks direct support for GPS data: to save battery, the Clip records raw signal strength data, which needs to be uploaded to company servers in order to be converted into actual coordinates.

The availability of contextual information such as location or even co-location of others (e.g., through social media records) greatly extends the ability of such images to be used as efficient retrieval cues, adding further information that might prompt an individual to recollect the actual event. An interesting yet highly contested source of such additional information is audio. Simple digital audio recorders (e.g., dictaphones) have long allowed continuous, unobtrusive recording of audio, yet their use is usually prohibited by law unless all parties recorded explicitly agree to such a recording. A recent crowdfunding effort, the wearable audio capture wristband "Kapture"[2] attempts to mitigate this by recording audio

[2]See `http://kaptureaudio.com/`

in a 60-second "audio loop". Only by visibly tapping the wristband one can save this last minute of audio in a persistent file, allowing the wearer to retroactively record the last 60 seconds of a discussion. By virtue of the very visible tap on the wristband, the designers hope to avoid the obvious social issue of covert recordings.

Recent technology makes it increasingly easy to record high-quality physiological data. Driven by the fitness and health industry, early technology like the Fitbit [3] records physical activity in the form of steps. The company's more recent products now also include stair climbing, sleep tracking, and continuous heart rate tracking. "Smart Watch" products from Apple, Microsoft, Samsung, LG and Motorola typically offer similar tracking features. Italy-based empatica claims that its E4 sensor wristband [4] offers a heart rate tracker that holds up to clinical trial standards. In addition, the E4 is able to record electrodermal activity (i.e., skin conductance), which allows one to measure arousal, stress, and excitement.

In short: today's technology offers up many tantalising possibilities for the continuous and comprehensive recording of many elements of our experiences, albeit not yet elements such as what we smell, touch and feel. These technologies make it possible to acquire a lifelog containing a wealth of potential cues to improve our memory of our past and our ability to remember future events. Not surprisingly, this also raises a number of important moral and societal questions. If lifelogging becomes commonplace and almost every aspect of our daily lives is recorded then what impact does this have on our own privacy and the privacy of others?

3.1. Privacy and Ethical Issues

Lifelogging poses significant privacy issues [33, 34, 35, 36]. Continuous image and video capture will inevitably capture private moments or personal information, e.g., when sitting in front of a computer screen or when entering a bathroom or bedroom. This might not be an issue if the data is never meant to be shared with others, though it raises the risk of accidental disclosure and opens up the possibility of data breaches. In 2014, several intimate pictures of celebrities were leaked after hackers broke into Apple's iCloud services[5] – a complete lifelog would be a treasure-trove for both nosy people and blackmailers. Gurrin remarks: "For example, if you're a lifelogger and you happen to walk by a school every morning, they could say you're taking pictures of children. It's so easy to pull out things that will actually cause you trouble" [36].

If sharing is done voluntarily, such records might not appear much different from today's social networking applications, where private pictures are often

[3] See `www.fitbit.com`

[4] See `www.empatica.com/e4-wristband`

[5] see `http://www.bbc.com/news/technology-29076899`.

shared. Given today's "social networking culture", it is not unlikely that lifelogging and social networking practices will intersect, making such sharing much more likely [37]. Since people freely share such images already, privacy concerns at first seem irrelevant. However, lifelogging will in most cases capture bystanders, friends, and family members alike, disclosing their whereabouts, activities, and appearances [38]. It is unlikely that a lifelogger would have permission from everybody captured on their lifelogging device, and even if so, such permission would practically amount to a "blank cheque" regarding the use and dissemination of such images. While many such pictures might be taken in public, where one would assume that no permissions are necessary, several jurisdictions require consent for taking one's picture even in public settings (e.g., Greece), thus effectively banning lifelogging in public [35]. Gurrin et al. [39] suggest using face-blurring technology, like Google's street view, to filter all access to captured imagery. Known contacts of the lifelogger could explicitly agree to appear in lifelogs by providing a set of images for a whitelist of faces that should not be removed upon access. Gurrin et al. suggest that the system store the unblurred images, only applying face blurring upon images access, allowing retroactive access to a person's image. This would, however, in principle allow hackers to circumvent this process entirely and thus access the non-blurred images in full.

While not something necessarily couched in terms of privacy, having one's own life memory may also pose significant problems for people with mental issues. Allen [33] calls this a "psychological hazard" and expects that this technology will "enable excessive rumination by persons experiencing unipolar or bipolar depression." Even "normal" people may have difficulties in "moulding and changing" their identity if they have access to a detailed daily documentary of their past.

One potential impact of lifelogging technology springs from its interaction on a societal scale. Lifelogging, in particular when stored on cloud services, holds the potential for state-level surveillance [33]. It is not hard to imagine law enforcement agencies considering lifelogs as fair game for hunting down terrorists, child molesters, or organised crime. While many may tolerate such activities in exchange for a perceived increase in personal safety, feature creep might extend the use of such data to fight tax evasion, social services fraud, speeding, and disputed insurance cases. The widespread collection of lifelogs may also accelerate the so-called "reversal of the burden of proof" [40], where "voluntary" lifelogs must be supplied to law enforcement in order to prove one's innocence [33].

The ability of lifelogs to provide a "record of truth" of someone's life, in particular when multiple streams intersect (images, GPS markers, iCal entries with location, geo-verified IP address in email sent at the time, etc), seems often to be universally accepted. However, Marx [41] points out the fallacy of believing that "data means knowledge", or that "more data means more knowledge": even very detailed records can leave significant room for interpretation, making it difficult, if not impossible, to prove intent or even actions from such records. As Taylor Owen from Columbia's TOW Center for digital journalism remarks [42]: "data does not equal fact" and "context matters enormously". Grudin [43] points out that the decision what to record and what not to record "profoundly changes the nature of that information and the contexts created from it. . . Anything that does not 'make the cut,' (i.e., is not captured) is invisible to someone inspecting the digital record at a different location or time. Anything that is recorded instantly achieves a potential pervasiveness and immortality that it did not have before." If lifelogs become commonplace, the assumption that they represent "ground truth" might need to be actively discouraged on a societal scale.

Few proposals exist on how to address lifelogging privacy issues. Gurrin et al. [39] and others have focused on face blurring, yet this technology is not completely reliable and might miss a large set of contextual data (e.g., Bluetooth traces, calendar entries involving attendee lists) that would still disclose the presence or activities of others. Dodge and Kitchin [1] suggest purposely designing "imperfection, loss and error" into lifelogging applications, though this hardly seems realistic – who would want to buy such a system? However, O'Hara [37] points out that most likely, a "perfect record" of daily life is anyway unrealistic: "Loggers might periodically forget to take GPS devices with them on their perambulations. New interests or a change of circumstance might result in more or less information being stored. The picture is likely to be patchy; undiscriminating information gathering does not entail an unblinking gaze." This resonates with Dourish and Bell's observation that much of the research in the area of wearable and ubiquitous computing tends to ignore what they call "the messiness of every day practice," [44] assuming instead that everything will work seamlessly in the future.

3.2. Lifelogging as a tool for memory augmentation

Researchers have investigated the use of SenseCam video data to help people with memory difficulties recollect events. Pauly-Takacs et al [45] applied this idea to assist a 13 year-old boy with profound episodic memory difficulties in remembering the details of a walk. They concluded that the lifelogging data was able to improve the boy's ability to look back and reflect upon the events and surmised that this was of tangible benefit to his rehabilitation. Wherton et al. [46] considered problems people with dementia have with typical tasks performed in a kitchen and how such patents could use videos from a lifelogging device. They conclude that the technology ameliorated the difficulties by prompting users when they had difficulties, thus fostering independence and quality of life. Berry et al. [21] used SenseCam to create pictorial diaries of a woman suffering from severe memory impairment and showed that by frequently reviewing past memories through these diaries she was able to recall approximately 80% of recent, experienced events.

Later work [47] showed more longitudinal effects of reviewing SenseCam images, demonstrating that they were twice as effective as detailed written accounts. Woodberry et al [22] showed that the structured review of SenseCam images taken by amnesics, aided by their spouses, resulted in superior recall of more significant details of events than a written diary. The SenseCam advantage was even greater after 1 and 3 months.

Iwamura et al. [48] considered the embarrassing problem of being unable to remember information about the person you are speaking to and devised and tested a system to help mitigate this problem. They focused on solutions that are able to rapidly call up previous instances of meetings between the user and the interlocutor, showing video clips of previous encounters and the locations and times of these. Other work [49] approached the question of how lifelogging can support human memory from a more theoretical perspective by considering memory problems mentioned in psychology literature and commenting on how technology could provide support mechanisms in these instances. They conclude that, once we understand how memory is used and also how it fails, technology could be used to help people recall forgotten information but also note that they might be better for confirming partially-remembered events.

Most research on the topic up until now has either focused on people with memory impairments or on users manually browsing through captured data, often with a fairly nebulous task at hand, for example determining whether or not the recorded data helped the user to remember being in a situation or at an event. Sellen et al. [20] studied how well automatically-captured images served as cues to memories about past events and found them to be equally as potent as those that were intentionally taken by users. In the relatively small-scale experiment nineteen participants wore SenseCam on two consecutive days and then were tested three or ten days later (and for a subset of the participants after four months). During the test session, recall was tested prior to and after seeing a subset of viewed images, which were used to test recognition memory. Critically, following a review of some images, participants showed superior free recall of additional events from that day than to an equated control day. The authors do, however, also warn that the results of their study suggest that these cues are only helpful in the short-term and may not be quite so influential in the long- term.

While these studies may have been useful to gain a cursory understanding of whether or not lifelogging data is useful for augmenting memory, they focus more on reminiscing than retrieval of truly useful and specific information or do not consider memory problems exhibited by healthy individuals. In order to investigate how such data might be used to ameliorate memory difficulties in everyday life we must be able to provide the user with more information about the recorded memories, in order to facilitate recall of *semantic* content. We now consider how the raw data from an image-based lifelogging tool can be stored, processed and presented so that this kind of advanced usage could be made possible.

4. From raw data to useful information

When considering the process of a computer system "memorising" lifelogging data, and particularly *episodic* memories, we can draw parallels with how the human mind goes about the same task. Since relatively few studies exist in computer science literature regarding what recordable information is useful for jogging memories, we instead look towards the field of psychology for clues. Psychology literature shows that there are three main phases associated with memory: encoding, storage and retrieval [50]. Each of these stages in the memorisation process can be investigated and used directly to instruct the design of memory aiding tools.

4.1. Segmenting life into episodes

In the first phase informative representations of each new memory are created based on the raw data obtained from the senses. However, before doing this the constant stream of memory representing one's experiences is segmented in the brain into meaningful atomic events or *episodes*. Individual episodes could be quite lengthy, for example having breakfast in the morning, attending a meeting at work or watching a football match. However they could also be quite momentary, for example seeing a bright flash of light or catching sight of a friend on the opposite side of the street.

Regardless of their length or perceived importance, it is important that each episode can be seen as a single semantic "unit" of time and that these units can be ordered in time relative to each other [51] and often also spatially. They are personal, auto-biographical events and do not necessarily relate to any other extra-episodic entity or reference and can be described (and indeed differentiated from each other) by their individual attributes. It is not known exactly to which level of granularity the mind separates memory into individual episodes, however it does seem that this varies depending on the event being encoded [52].

In much the same way that a search system will take you directly to the web page or article you are looking for, and not just the main page of the site it is found on, lifelogging data should also be intelligently separated into smaller constituent parts. Early approaches to this separated the media into segments of pre-defined length [53], however this is far too simplistic as real events are not always of a single fixed duration. In the field of video retrieval techniques have been developed to detect boundaries between events, often involving the calculation of pixel intensities, histograms of those intensities and edge detection. Boundaries in the video sequences are then determined to be where the difference between successive frames exceeds a given threshold [54]. While such methods can be reliable for video, they cannot be so easily applied to lifelogging data as it usually has a much lower frame-rate, often less than a single frame per second.

Doherty et al. [55] found image-based approaches to event detection for SenseCam images to be quite reliable, however the methods tested tended to struggle with many common occurrences, such as someone running to catch a bus as successive images are highly unstable and very dissimilar. However, we would clearly prefer to assign the whole period of running to a single event, rather than separating it into many very short ones. Another problematic situation is when the wearer is performing essentially one task, such as talking to a friend, but occasionally turns to glance at other events, such as someone entering the room. Again, in this instance even a single frame captured during the wayward glance would result in the single event of chatting to a friend being partitioned into multiple episodes.

To mitigate this issue, Doherty et al. use an approach based on Hearst's Text Tiling algorithm [56] where instead of comparing pairs of single adjacent frames, blocks of adjacent frames are compared. They use a block size of 5 images and slide forward along the sequence of all images, making comparisons at each step. The authors also incorporate data from other sources such as the additional sensors available on the SenseCam which includes temperature sensors, accelerometers and passive-infrared sensors. The values from all of these sensors can be fused with data from image boundary detection to create a more nuanced algorithm far less prone to making mistakes like those described in the above examples.

The authors conclude that the Text Tiling technique works well for this problem and that incorporating other sensor information significantly improves performance, going on to claim that "segmentation of a lifelog of images into events [is] a solved problem." However, if lifelogging data is to be used to augment human memory, these sophisticated methods may still not be sufficiently powerful as even a small number of errors may be problematic. Furthermore, there may be cases where this fixed discretisation will not be appropriate as the user may instead want a summary of a number of related events or may wish to remember only a single element of an event, such as a logo they happened to glance at. This will require more powerful and sophisticated methods of analysing the data which allow for flexibility in the definition of a time "unit."

4.2. Encoding with context

Assuming the data has been segmented appropriately into discrete units, each individual episode must then be committed to memory so that it can be later refound. While this process may be conducted with conscious effort, the majority of what we encode is done "incidentally," merely a by-product of processing. However consciously it is done, the mind fits the new memory into its existing framework of memories by way of association, using points of reference to anchor the new memory to existing ones or to other key elements in memory, often referred to as context [17]. Very often one is unable to say at exactly what time of day a specific remembered event took place or exactly where, but multiple events can usually be placed in time and space relative to each other [57]. Consider how often a useful tactic, when you have forgotten something, is to think of where you were when you heard about it or who else was present at the time. This often works because these points of reference anchor the memory, underpinning it within the mind and guiding the mind towards the specific desired piece of information.

Studies have shown that the degree to which new information can be related to information already stored in long-term memory is directly associated with how well it will be stored and, subsequently, how easily it can be re-found [57]. The level to which a new memory is tied to existing ones is dependent on several factors including how much attention the person pays to the stimuli at the time of encoding and whether or not the person encoding the memory physically performed an action or was simply told or read about it [17]. Even a very fragmentary, incomplete set of contextual features from a half-remembered encounter can be sufficient to jog the brain into retrieving a complete and accurate account of the event [58]. The key is to have enough context attached to each memory at the time of encoding to allow any episode to be recalled based on whatever information is available to the person trying to remember, even if it is somewhat scant.

The encoding of memories can therefore be described as set of features containing: some kind of factual representation of the event, the set of contextual ties - a characterisation of the spatiotemporal context in which it occurred - and a reference to the person him/herself [59]. The importance of this spatiotemporal context can be demonstrated by the fact that being in the same room as you were when the event you are trying to recall took place can significantly increase the likelihood of you remembering it accurately [60]. Evidence has been shown for visual, spatial, acoustic, temporal and self encoding (binding the memory to the person who experienced it). This means that the mind uses a variety of different clues to narrow down memories and ensure that it is selecting the relevant one which is achieved by considering memories relative to each other and not atomically.

As discussed earlier, the likelihood of a searched-for memory being successfully recalled is dependent on how it was stored at the time of encoding and not on how alert or attentive the person is at the time of attempted recall [61]. What is clear is that the proper and rigorous encoding of the memory at the time it is experienced is highly influential in its potential for future successful recall. As Tulving puts it: "specific encoding operations performed on what is perceived determine what is stored, and what is stored determines what retrieval cues are effective in providing access to what is stored." This being the case, it is clear that to be successful in its task a memory aid would have to provide users with similar clues to those employed by the mind.

We consider a future system in which data captured by a lifelogging device would be regularly uploaded to some form of server which would subsequently perform any necessary processing and store the resulting episodes in a database for future retrieval. Each episode could then be "tagged" with metadata describing the context in which it was captured, its relationship to other captured episodes and references to any other relevant items within the database. In order to ensure that all tags are kept together with their corresponding images, it may also be sensible to embed the meta information into the image files themselves using EXIF tags or any similar scheme [62]. The user would be at liberty to assign their own tags to the data stored, however due to the sheer amount of data recorded [63] it would certainly not be reasonable to expect the user to do this for all of the recorded information. Much like the way the human mind processes memories from raw data from the senses, the data from the lifelogging device would have to be automatically processed and useful tags automatically assigned.

Many pieces of contextual information, such as GPS location, temperature and time can be trivially captured from the device and, with little processing, directly entered into the database. In addition to these, items posted by the user and their friends on social media websites (such as Facebook and Twitter) at the time of each episode could also be included as meta-data. Since people often post about what they are doing or thinking, these may serve as particularly useful and easy to obtain sources of contextual information. For a number of other potentially useful contextual features further processing will be required as these details cannot simply be obtained from the device's recordings or directly from the Internet. The ways in which these and other contextual features might be utilised to allow users to search though their memories will be described in the next section.

4.2.1. Using people to jog memories

Research into refinding of emails has shown that people often act as "hubs" or "reference points" between other people and objects, assisting users to locate the emails they are looking for [64]. In many instances a user was unable to accurately recollect the subject of an email but was nevertheless able to find it by remembering who it was sent from or to. The suitability of video data to remind people of others they had spoken to was investigated in recent work by Iwamura et al. [48]. However, to make the technology truly useful it would be necessary to identify people within data captured by a lifelogging device and subsequently tag the episodes in which they feature with a reference to each person's details. The system could then link this to any other stored data related to them, for example emails received from them, other photos and videos featuring them, calendar events they are attending or have attended or their social media accounts.

There has been significant interest in recent years in the automatic recognition of human faces in video streams and still images with software such as iPhoto and services such as Google images and Facebook implementing such technology. It is worth noting that for a human being these tasks seem to be fairly simple since a healthy human brain is able to perform both functions with levels of reliability approaching 100%. It is not yet known exactly how the brain goes about these tasks, although the clear aptitude of it in this case has led many neuroscientists to suggest that it may be a dedicated process - i.e. that there might be an area of the brain that has evolved specifically for these problems [65] - although there are still instances (such as under unusual lighting conditions) where the brain struggles. Unfortunately, this apparent ease of processing is certainly not the case for machine learning algorithms, although a large part of this is likely due to the fact that normal images are 2-dimensional and as such lack any of the depth features available to the human sensory system.

To recognise people in an image, the system must first detect any potential human faces present. Each candidate face must be isolated and discriminated from the image's background, a problem made non-trivial by the dynamic and varied nature of human faces as well as variations caused by different lighting conditions or angles of the shot [66]. While early approaches tried to mimic the assumed mechanisms of the brain by using edge detection to look for face-like oval shapes [67], modern face detection algorithms instead rely on machine learning technologies to learn from large data sources. These use a combination of edge detection and low-level pattern-based features together with sophisticated multi-feature classification methods and do not require a set of pre-defined facial feature templates to be defined [68, 69]. Each of these candidate faces must then be analysed for features, allowing the system to match them with a pre-existing database of known people. For a comprehensive overview of developments in this area the reader is encouraged to read a literature review of the field by Tolba et al. [70].

A large-scale evaluation of the most competitive face recognition methods was conducted in 2002, testing 10 different approaches on 121,589 images of 37,437 people [71]. To be reflective of real-world tasks, the images were taken from the U.S. Department of State's Mexican non-immigrant Visa archive. The best-performing method was able to achieve a 78% success rate for females and 79% for males, although these rates decreased to slightly below 50% in the case of images taken outside, where natural light makes identification much more challenging. Considering the problem we are interested in here and the relatively low level of fidelity offered by cameras on lifelogging devices, it is encouraging that the results of these tests suggest performance is not adversely affected by the resolution of images used.

Even a cursory glance at the literature on the topic of facial recognition makes it clear that there is some progress yet to be made, however using this technology would allow episodes to be tagged with people in many instances. It is also worth considering that the sheer number of images taken by lifelogging devices would result in two key benefits. Firstly, it could supply any face detection and recognition system with large amounts of training data, although this would require the user to manually tag faces on a large sample of photos. This may not be as unlikely as it appears at first glance since many people tag themselves and friends on photos posted to social media which could be used as a initial source of such data, should the user be happy to grant access to their profile. Secondly, and perhaps more crucially, it means that even over the period of a short half an hour meeting upwards of 30 face-on images of the interlocutor could be taken, giving the recognition algorithm many possibilities to obtain an exact match. Even if exact matches are not made, the probability of the correct identity being present in the images is likely to be high when combined over the series images taken.

4.2.2. Identifying objects and higher-level concepts

An obvious extension to detecting people would be to also detect other objects in the images as these may also serve as useful hooks to jog a user's memory. For example you may not remember when or where an event of interest took place but may be able to recall that there was a fountain nearby or particular brand of clothing store. The field of object detection in images is still in its infancy at the time of writing. However, as with facial recognition, we can look to the domain of videos and computer vision for appropriate literature. The problem of concept detection can be defined as attempting to assign probability values to the presence or absence of a pre-defined set of objects within a video clip or image [72].

Many approaches to this problem use a variety of low-level features of the image (e.g. colours, textures and shapes) which are then compared to higher-level representations. Using a variety of - typically Bayesian probabilistic - methods, fuzzy assignments are made to a database of known objects (or "concepts") via the high-dimensional feature space and through this probabilities are derived [73, 74]. More recently a Bag-of-Visual words approach, where areas of images with common visual features, known as "patches", are used as prototypical instances to describe images. These are typically more discriminative than more rudimentary features like colours and textures and are often described by their deviation from a "universal" patch model [75]. Recent developments in machine learning technology [76] make use of deep-learning neural networks have resulted in ever better classification results, including applications for object recognition [77].

In the area of lifelogging, Byrne et al. [72] attempted to adapt these methods to detect objects from low-fidelity images obtained from SenseCams. Manual analysis of SenseCam images identified a set of 27 "core concepts", defined to be those that are collection-independent, such as hands, buildings, the sky and meetings. Note that these concepts are not just restricted to physical objects but can also represent less tangible notions such as "shopping" and "holding a mobile phone." In total 51,396 training examples are obtained in this manner with counts per concept ranging from nearly 20,000 to as few as 3. Each training image is divided into a number of overlapping rectangular crops and two sets of standard image features are extracted which model surface textures [74]. Support Vector Machines (SVM) are used to train a classifier based on the training data and this classifier is then used to identify the existence of the pre-defined concepts in test images.

The performance of the approach was evaluated compared to manual judgements of the system's output by 9 human annotators. The accuracy of the method in terms of identifying the existence of concepts within images ranged from 82 to 23%, depending on the concept. There was a strong linear correlation between the prediction accuracy and the number of training images for each concept, suggesting that results would improve if the system were given more labelled data to work from. The approach faired much better when attempting to determine the absence of concepts within images with negative accuracy ranging from 100 to 45%. Overall performance, considering both positive and negative examples over all concepts, was 75%.

4.3. Retrieving memories

The final phase of remembering occurs when we wish to recall a past event or memory [78]. The brain attempts to use any and all pieces of context available to isolate the specific memory of interest; according to Tulving [52]: "some sort of a more complex interaction between stored information and certain features of the retrieval environment seems to be involved in converting a potential memory into conscious awareness of the original event." The brain has no identifiable interface for this task - we simply concentrate on things related to the memory we wish to retrieve and it either surfaces or does not. For any kind of lifelogging-based memory aid to be useful, however, it is crucial to consider how the user of such a system can be supported when searching and browsing through recorded memories.

Assuming a person sleeps for 8 hours per day, uses a lifelogger when awake and the devices captures an image every half a minute, in only a single day 1,920 images will be captured. Even if the person were to very rapidly browse through these images at a rate of 3 per second, it would take 16 minutes each day to go through them all. While lifelogging devices may allow for automatic capture of a person's day, this still does not resolve the problem of trying to augment human memory. For this raw data to be useful, even after it has been well processed and tagged, there must be a useful and intuitive interface to allow the user access to the potentially very small parts of the whole collection that is needed to jog the memory of a particular event.

For insights on how best to design such an interface we can consider the literature from the field of Human Computer Interaction in the area of image browsing and search systems. As the number of images captured by digital devices has increased, the necessity for an interface with which to manage and catalogue them has increased accordingly [79]. Popular software such as Apple's iPhoto[6] and Google's Picasa[7] allow users to group images into albums and search via tags they can explicitly assign. These application also attempt to identify faces within photos and assign them to know people, allowing searches to by made by person.

Many interfaces are based around time [80, 81] - allowing users to find images based on when they were taken - and location (often ascertained from GPS data) [82] - grouping images by where they were taken. These features, together with tagging, allow searches through the image collection to be more precise and discriminative and to some extent take advantage of the mental encoding procedure of the mind described earlier. Chen et al. [81] describe time and location as being "the main factors in episodic memory" and compare a photo browsing interface which clusters images based on time and place with 4 more "standard" interfaces. Use of the spatio-temporal browser resulted in a significant reduction in the time require to complete a selection of photo refinding tasks and users gave the interface significantly higher satisfaction scores. [83]

Elsweiler et al. [14] conducted a diary study to identify common everyday memory problems to inform the design of a photo browsing tool. They found that people often remember and refine memories through what they describe as "retrieval journeys" where they initially search broadly using a small piece of contextual information and then browse through this subset of photos which may in turn trigger more details about the desired photo to surface. They designed 3 photo browsing tools, one of which ignored normal photo browsing tool features (such as albums) and was instead designed to support these journeys, and tested these out in a second study. Their results showed that although users often remembered a number of contextual details of the photo they wanted to find, with more traditional interfaces they tended only to use these sparingly and in isolation. Users preferred the novel interface, employing more forms of context and were generally more pleased with the results of their searches.

Interfaces for SenseCam data have been developed [31, 84] which employ some of the automatic encoding steps outlined earlier in this paper: breaking up the stream of images into episodes; selecting key frames to represent each episode and tagging each image with contextual data from the device. Rather than simply displaying all of the captured images in a list or as thumbnails, which may be overwhelming, Lee et al. [31] algorithmically select the 20 most significant or unique events from each day and choose, from each of these, a "representative frame" (see figure 2). After experimentation with more sophisticated methods, each representative frame is chosen by simply taking the middle frame from each event.

[6]iPhoto official web site - https://www.apple.com/mac/iphoto/
[7]Google Picasa official web site - http://picasa.google.com

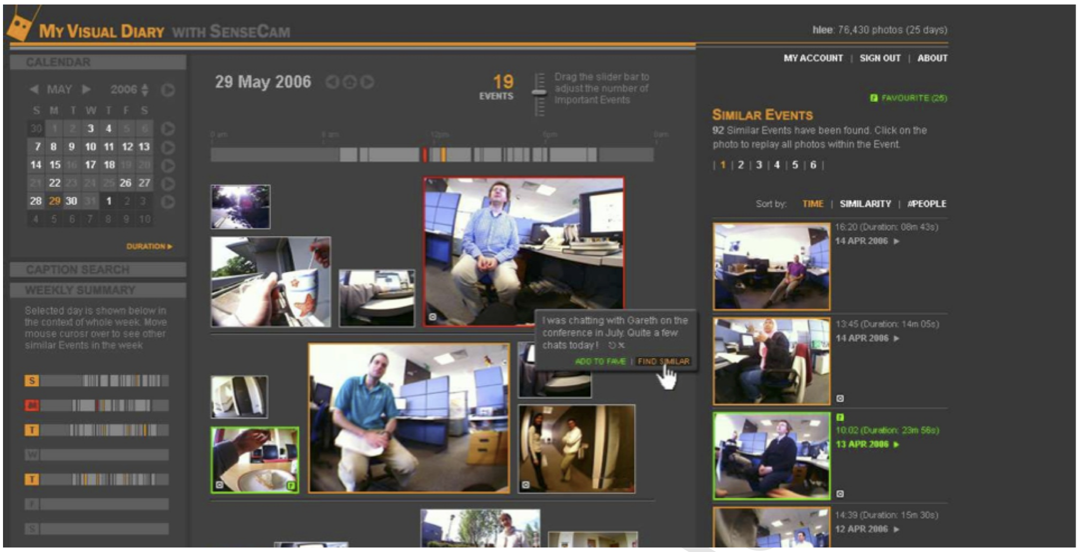

Figure 2: Screenshot of a recorded events browsing tool (developed for SenseCam [31]).

The more "novel" the event, the larger the representative image will be made, attracting the attention of the user to more interesting events. The novelty or uniqueness of an event is determined by how different it is from the average event in the same week. Hovering over an event will cycle through the images pertaining to it and highlight the window of time it represents in a timeline which appears above the event thumbnails. Using a collection of 2.5 years' worth of SenseCam images, Doherty et al. [85] showed that allowing the user to refine searches, using computed meta-data such as who is present in each event, increased the speed and accuracy of recall.

While these examples provide useful insights into how to manage the images captured during lifelogging, they do not address the storage and organisation of other forms of data collected. MyLifeBits, part of the SenseCam project, represents an early prototype storage system for various forms of collected data, including: images, video, emails, sensor readings and health-related information [79]. Originally, the project relied on meticulous naming of files and prudent folder use, however as the number of sources (and amount of captured data) grew, this schema became increasingly unsuitable. The limitations of using the desktop file and folder metaphor were obvious and as such the project turned to database technology, allowing metadata to be readily assigned to items and complex searches queries to be quickly and efficiently posed to the system.

The system was designed around the idea of search being the most crucial component, allowing items to be linked together via their respective metadata, corresponding to how recall works in the brain. Items related to a given query can be displayed as either a list, a set of thumbnails, in a timeline view or on a map (if GPS data is available) and may be annotated with voice or text which is added to that file's metadata store. Any item may be associated with any other, for example images can be annotated with the people present in them. Furthermore, items are split between the user's personal and professional lives and between current - ongoing projects or recent appointments - and archived items - events in the past, former relationships and collaborations.

Having used the system personally for some time, the authors noted that it is "freeing, uplifting" to have a "surrogate memory" available and found that the more information captured in the system, the more useful it becomes as more correlations between the data emerge [79]. They mention a feeling of security that comes from knowing that the data is being collected automatically and continuously and were frustrated by instances where data they were interacting with could not be stored due to security or copyright reasons. The need for users to manage and curate, in some fashion, their data collections is mentioned and that this might be too onerous a task, leading to the suggestion that automated assignment of metadata tags would be very helpful. In terms of organisation, it is suggested that some combination of a tagging-based and hierarchical system would be most suitable, although this may not be well adopted as people may be conditioned by existing paradigms used by existing email clients and file systems.

5. Conclusions

Lifelogging technology allows us to capture digital snapshots representing every moment of our lives and to store this information for posterity, presenting clear opportunities for augmenting human memory. Unfortunately the sheer amount of information captured, and the form this takes (namely images and video), brings problems of how to effectively catalogue and manage this information. The question of how to effectively use this information to assist people in remembering episodes from the past requires a thorough understanding of how the mind itself deals with recall. In this article we have composed key insights and results from work from different fields with the aim of providing an overview of the current state of knowledge around this research topic. This overview shows that much of the technologies and understanding necessary to develop a human memory augmentation solution exists, however they currently exist separately and have not yet been combined into a single, workable and efficient solution. While early studies have shown promise for lifelogging technology as an aid for those with chronic memory problems, the technology could also be useful for those without such problems.

Simply integrating existing techniques into a single tool would not result in a completely usable solution which would make memory loss a thing of the past. However, the necessary early steps toward such a system have been made and in future work it will be possible to unite these into a common framework. A prototype toolchain could be developed with data recorded by a lifelogging device which would be uploaded regularly (perhaps even automatically) to a server which would process the data. The server would first determine which images should be kept and which should be discarded. The remaining images would be segmented into episodes and then analysed to produce useful metadata which would then be encoded into the image via an EXIF-like scheme. Improvements could be made to the work of Byrne et al. [72] by using contemporary image analysis and recognition tools [77] allowing episodes to be better tagged automatically.

To facilitate rapid searching and browsing of the data collected, a database would be maintained simultaneously and constantly kept in sync with the state of the collection. Methods of displaying this wealth of data to the user - and whether or not to do so pro-actively - will require further studies on how people explore and search lifelogs. Regular reviews of key memories may help the user better recall events during instances of spontaneous recall and when the prosthesis device is not available. Further development dictated by the results of studies performed using this toolchain would provide researchers with a means to investigate how such a system might be best utilised and where further improvements may be necessary.

Beyond technical issues of how best to use such a system and how to develop user interfaces to work in concert with the mind to navigate and find hard-to-remember episodes, longitudinal studies could investigate how use of the system alters the behaviour of its users. Could, for example, the user's own memory be improved by subtle and well-timed prompting and reviewing of key memories? Could long-term use of the device actually have a detrimental effect on the user's ability to remember naturally (i.e. when not relying on the digital surrogate)? Would users sometimes completely forget an episode and be surprised by viewing it in their personal memory collection or would they remember the event very differently from how it actually transpired? As highlighted earlier, such technology brings about new ethical questions. How can we best go about exploiting this technology without it having a detrimental effect on the privacy and rights of individuals and where should a compromise be drawn? Finally, could we use such technology together with research on "recall-induced forgetting" to assist people in forgetting unwanted memories and would it be truly ethical to do so?

Acknowledgements Thanks go to Ana Javornik for her help in proof-reading the article and for her very useful comments on the content.

References

[1] M. Dodge and R. Kitchin, "" outlines of a world coming into existence": Pervasive computing and the ethics of forgetting," *Environment and Planning B*, vol. 34, no. 3, pp. 431–445, 2007.

[2] C. Gurrin, A. F. Smeaton, and A. R. Doherty, "Lifelogging: Personal big data," *Foundations and Trends in Information Retrieval*, vol. 8, no. 1, pp. 1–125, 2014.

[3] M. Lamming, P. Brown, K. Carter, M. Eldridge, M. Flynn, G. Louie, P. Robinson, A. Sellen, M. Lamming, P. Brown, K. Carter, M. Eldridge, M. Flynn, G. Louie, P. Robinson, and A. Sellen, "The design of a human memory prosthesis," *The Computer Journal*, vol. 37, pp. 153–163, 1994.

[4] A. M. Surprenant and I. Neath, *Principles of memory.* Psychology Press, 2013.

[5] N. J. Slamecka, "Proactive inhibition of connected discourse.," *Journal of experimental psychology*, vol. 62, no. 3, p. 295, 1961.

[6] G. D. Brown, I. Neath, and N. Chater, "A temporal ratio model of memory.," *Psychological review*, vol. 114, no. 3, p. 539, 2007.

[7] L. Goddard, L. Pring, and N. Felmingham, "The effects of cue modality on the quality of personal memories retrieved," *Memory*, vol. 13, no. 1, pp. 79–86, 2005.

[8] J. R. Anderson and L. M. Reder, "The fan effect: New results and new theories.," *Journal of Experimental Psychology: General*, vol. 128, no. 2, p. 186, 1999.

[9] R. N. Shepard, "Recognition memory for words, sentences, and pictures," *Journal of verbal Learning and verbal Behavior*, vol. 6, no. 1, pp. 156–163, 1967.

[10] H. L. Roediger and J. D. Karpicke, "Test-enhanced learning taking memory tests improves long-term retention," *Psychological science*, vol. 17, no. 3, pp. 249–255, 2006.

[11] M. C. Anderson, R. A. Bjork, and E. L. Bjork, "Remembering can cause forgetting: retrieval dynamics in long-term memory.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 20, no. 5, p. 1063, 1994.

[12] W. S. Terry, "Everyday forgetting: Data from a diary study," *Psychological reports*, vol. 62, no. 1, pp. 299–303, 1988.

[13] M. Eldridge, A. Sellen, and D. Bedkerian, *Memory Problems at Work: Their Range, Frequency and Severity.* Rank Xerox, EuroPARC, 1992.

[14] D. Elsweiler, I. Ruthven, and C. Jones, "Towards memory supporting personal information management tools," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 7, pp. 924–946, 2007.

[15] N. Unsworth, B. D. McMillan, G. A. Brewer, and G. J. Spillers, "Individual differences in everyday retrospective memory failures," *Journal of Applied Research in Memory and Cognition*, vol. 2, no. 1, pp. 7–13, 2013.

[16] F. I. Craik, R. Govoni, M. Naveh-Benjamin, and N. D. Anderson, "The effects of divided attention on encoding and retrieval processes in human memory.," *Journal of Experimental Psychology: General*, vol. 125, no. 2, p. 159, 1996.

[17] F. I. Craik and R. S. Lockhart, "Levels of processing: A framework for memory research," *Journal of verbal learning and verbal behavior*, vol. 11, no. 6, pp. 671–684, 1972.

[18] W. Koutstaal, D. L. Schacter, M. K. Johnson, and L. Galluccio, "Facilitation and impairment of event memory produced by photograph review," *Memory & Cognition*, vol. 27, no. 3, pp. 478–493, 1999.

[19] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, "Sensecam: A retrospective memory aid," in *Proceedings of the 8th International Conference of Ubiquitous Computing (UbiComp 2006)*, pp. 177–193, Springer Verlag, September 2006.

[20] A. J. Sellen, A. Fogg, M. Aitken, S. Hodges, C. Rother, and K. Wood, "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*, CHI '07, (New York, NY, USA), pp. 81–90, ACM, 2007.

[21] E. Berry, N. Kapur, L. Williams, S. Hodges, P. Watson, G. Smyth, J. Srinivasan, R. Smith, B. Wilson, and K. Wood, "The use of a wearable camera, sensecam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report," *Neuropsychological Rehabilitation*, vol. 17, no. 4-5, pp. 582–601, 2007.

[22] E. Woodberry, G. Browne, S. Hodges, P. Watson, N. Kapur, and K. Woodberry, "The use of a wearable camera improves autobiographical memory in patients with alzheimer's disease," *Memory*, vol. 23, no. 3, pp. 340–349, 2015.

[23] S. Hodges, E. Berry, and K. Wood, "Sensecam: A wearable camera that stimulates and rehabilitates autobiographical memory," *Memory*, vol. 19, no. 7, pp. 685–696, 2011.

[24] L. Tan and G. Ward, "A recency-based account of the primacy effect in free recall.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 26, no. 6, p. 1589, 2000.

[25] V. Bush, "As we may think," *SIGPC Note.*, vol. 1, pp. 36–44, Apr. 1979.

[26] S. Mann, ""smart clothing": Wearable multimedia computing and "personal imaging" to restore the technological balance between people and their environments," in *Proceedings of the Fourth ACM International Conference on Multimedia*, MULTIMEDIA '96, (New York, NY, USA), pp. 163–174, ACM, 1996.

[27] S. Mann, "Wearable computing: A first step toward personal imaging," *Computer*, vol. 30, pp. 25–32, Feb. 1997.

[28] S. Mann, J. Fung, C. Aimone, A. Sehgal, and D. Chen, "Designing eyetap digital eyeglasses for continuous lifelong capture and sharing of personal experiences," *Alt. Chi, Proc. CHI 2005*, 2005.

[29] G. Hayes, S. Patel, K. Truong, G. Iachello, J. Kientz, R. Farmer, and G. Abowd, “The personal audio loop: Designing a ubiquitous audio-based memory aid,” in *Mobile Human-Computer Interaction - MobileHCI 2004*, vol. 3160, pp. 168–179, 2004.

[30] M. A. Conway, “Episodic memories,” *Neuropsychologia,*, vol. 47, no. 11, pp. 2305–2313, 2009.

[31] H. Lee, A. Smeaton, N. O’Connor, G. Jones, M. Blighe, D. Byrne, A. Doherty, and C. Gurrin, “Constructing a sensecam visual diary as a media process,” *Multimedia Systems*, vol. 14, no. 6, pp. 341–349, 2008.

[32] K. Wolf, A. Schmidt, A. Bexheti, and M. Langheinrich, “Lifelogging: You’re wearing a camera?,” *IEEE Pervasive Computing*, vol. 13, no. 3, pp. 8–12, 2014.

[33] A. L. Allen, “Dredging up the past: Lifelogging, memory, and surveillance,” *The University of Chicago Law Review*, pp. 47–74, 2008.

[34] I. Askoxylakis, I. Brown, P. Dickman, M. Friedewald, K. Irion, E. Kosta, Langheinrich, *et al.*, “To log or not to log?-risks and benefits of emerging life-logging applications,” 2011.

[35] N. Davies, A. Friday, S. Clinch, C. Sas, M. Langheinrich, G. Ward, and A. Schmidt, “Security and privacy implications of pervasive memory augmentation,” *IEEE Pervasive Computing*, vol. 2015, 2015.

[36] C. Gurrin, “Lifelogging: The end of privacy,” October 2014.

[37] K. OHara, M. M. Tuffield, and N. Shadbolt, “Lifelogging: Privacy and empowerment with memories for life,” *Identity in the Information Society*, vol. 1, no. 1, pp. 155–172, 2008.

[38] D. H. Nguyen, G. Marcu, G. R. Hayes, K. N. Truong, J. Scott, M. Langheinrich, and C. Roduner, “Encountering sensecam: personal recording technologies in everyday life,” in *Proceedings of the 11th international conference on Ubiquitous computing*, pp. 165–174, ACM, 2009.

[39] C. Gurrin, R. Albatal, H. Joho, and K. Ishii, “A privacy by design approach to lifelogging,” in *Digital Enlightenment Yearbook 2014: Social Networks and Social Machines, Surveillance and Empowerment* (K. O’Hara, M.-H. C. Nguyen, and P. Haynes, eds.), pp. 49–73, IOS Press, 2014.

[40] A. Galetta, “The changing nature of the presumption of innocence in todays surveillance societies: rewrite human rights or regulate the use of surveillance technologies?,” *European Journal of Law and Technology*, vol. 4, no. 2, 2013.

[41] G. Marx, “Some information age techno-fallacies,” *Journal of Contingencies and Crisis Management*, vol. 11, no. 1, pp. 25–31, 2003.

[42] T. Owen, “What the tesla affir tells us about data journalism.” TOW Center Blog, `http://towcenter.org/what-the-tesla-affair-tells-us-about-data-journalism/`, February 21 2013.

[43] J. Grudin, “Desituating action: Digital representation of context,” *Human-Computer Interaction*, vol. 16, no. 2, pp. 269–286, 2001.

[44] G. Bell and P. Dourish, “Yesterday’s tomorrows: notes on ubiquitous computing’s dominant vision,” *Personal and Ubiquitous Computing*, vol. 11, no. 2, pp. 133–143, 2007.

[45] K. Pauly-Takacs, C. J. A. Moulin, and E. J. Estlin, “Sensecam as a rehabilitation tool in a child with anterograde amnesia,” *Memory*, vol. 19, no. 7, pp. 705–712, 2011.

[46] J. P. Wherton and A. F. Monk, “Problems people with dementia have with kitchen tasks: The challenge for pervasive computing,” *Interacting with Computers*, vol. 22, no. 4, pp. 253 – 266, 2010.

[47] G. Browne, E. Berry, N. Kapur, S. Hodges, G. Smyth, P. Watson, and K. Wood, “Sensecam improves memory for recent events and quality of life in a patient with memory retrieval difficulties,” *Memory*, vol. 19, no. 7, pp. 713–722, 2011.

[48] M. Iwamura, K. Kunze, Y. Kato, Y. Utsumi, and K. Kise, “Haven’t we met before?: A realistic memory assistance system to remind you of the person in front of you,” in *Proceedings of the 5th Augmented Human International Conference*, AH ’14, (New York, NY, USA), pp. 32:1–32:4, ACM, 2014.

[49] Y. Chen and G. J. F. Jones, “Augmenting human memory using personal lifelogs,” in *Proceedings of the 1st Augmented Human International Conference*, AH ’10, (New York, NY, USA), pp. 24:1–24:9, ACM, 2010.

[50] M. W. Eysenck, *Principles of cognitive psychology.* Psychology Press, 2001.

[51] E. Tulving, “Episodic and semantic memory 1,” *Organization of Memory. London: Academic*, vol. 381, p. e402, 1972.

[52] E. Tulving and D. M. Thomson, “Encoding specificity and retrieval processes in episodic memory.,” *Psychological review*, vol. 80, no. 5, p. 352, 1973.

[53] Z. Wang, M. D. Hoffman, P. R. Cook, and K. Li, “Vferret: content-based similarity search tool for continuous archived video,” in *Proceedings of the 3rd ACM workshop on Continuous archival and retrival of personal experences*, pp. 19–26, ACM, 2006.

[54] H. Zhang, A. Kankanhalli, and S. W. Smoliar, “Automatic partitioning of full-motion video,” *Multimedia systems*, vol. 1, no. 1, pp. 10–28, 1993.

[55] A. Doherty and A. Smeaton, “Automatically segmenting lifelog data into events,” in *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS ’08*, pp. 20–23, May 2008.

[56] M. A. Hearst and C. Plaunt, “Subtopic structuring for full-length document access,” in *Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR ’93, (New York, NY, USA), pp. 59–68, ACM, 1993.

[57] C. P. Thompson, J. J. Skowronski, S. F. Larsen, and A. L. Betz, *Autobiographical memory: Remembering what and remembering when.* Psychology Press, 2013.

[58] W. A. Wagenaar, "My memory: A study of autobiographical memory over six years," *Cognitive psychology*, vol. 18, no. 2, pp. 225–252, 1986.

[59] G. M. Davies and D. M. Thomson, *Memory in context: Context in memory.* John Wiley & Sons, 1988.

[60] S. M. Smith, *Memory in context: Context in memory.*, ch. Environmental Context-Dependent Memory, pp. 13–34. John Wiley & Sons, 1988.

[61] A. P. Yonelinas, "The nature of recollection and familiarity: A review of 30 years of research," *Journal of memory and language*, vol. 46, no. 3, pp. 441–517, 2002.

[62] C.-J. Jang, J.-Y. Lee, J.-W. Lee, and H.-G. Cho, "Smart management system for digital photographs using temporal and spatial features with exif metadata," in *Digital Information Management, 2007. ICDIM'07. 2nd International Conference on*, vol. 1, pp. 110–115, IEEE, 2007.

[63] J. Farringdon and V. Oni, "Visual augmented memory (vam)," in *Proceedings of the 4th IEEE International Symposium on Wearable Computers*, pp. 167–, 2000.

[64] M. Harvey and D. Elsweiler, "Exploring query patterns in email search," in *Proceedings of the 34th European Conference on Advances in Information Retrieval*, pp. 25–36, Springer-Verlag, 2012.

[65] I. Gauthier, M. Behrmann, and M. J. Tarr, "Can face recognition really be dissociated from object recognition?," *J. Cognitive Neuroscience*, vol. 11, pp. 349–370, June 1999.

[66] V. Bruce, "Perceiving and recognising faces," *Mind & Language*, vol. 5, no. 4, pp. 342–364, 1990.

[67] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–741, 1995.

[68] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 696–706, 2002.

[69] C. Liu, "A bayesian discriminating features method for face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 725–740, June 2003.

[70] A. Tolba, A. El-Baz, and A. El-Harby, "Face recognition: A literature review," *International Journal of Signal Processing*, vol. 2, no. 2, pp. 88–103, 2006.

[71] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone, "Face recognition vendor test 2002," in *Analysis and Modeling of Faces and Gestures, 2003.*, p. 44, IEEE, 2003.

[72] D. Byrne, A. R. Doherty, C. G. Snoek, G. J. Jones, and A. F. Smeaton, "Everyday concept detection in visual lifelogs: Validation, relationships and trends," *Multimed. Tools Appl.*, vol. 49, pp. 119–144, Aug. 2010.

[73] J. C. Van Gemert, C. G. Snoek, C. J. Veenman, A. W. Smeulders, and J.-M. Geusebroek, "Comparing compact codebooks for visual categorization," *Computer Vision and Image Understanding*, vol. 114, no. 4, pp. 450–462, 2010.

[74] J. Krapac, J. Verbeek, and F. Jurie, "Modeling spatial layout with fisher vectors for image categorization," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 1487–1494, IEEE, 2011.

[75] J. Sánchez, F. Perronnin, T. Mensink, and J. Verbeek, "Image classification with the fisher vector: Theory and practice," *International journal of computer vision*, vol. 105, no. 3, pp. 222–245, 2013.

[76] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *ACM conference on Multimedia*, pp. 675–678, ACM, 2014.

[77] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," *arXiv preprint arXiv:1310.1531*, 2013.

[78] G. Gillund and R. M. Shiffrin, "A retrieval model for both recognition and recall.," *Psychological review*, vol. 91, no. 1, p. 1, 1984.

[79] J. Gemmell, G. Bell, and R. Lueder, "Mylifebits: a personal database for everything," *Communications of the ACM (CACM)*, vol. 49, pp. 88–95, January 2006. also as MSR-TR-2006-23.

[80] M. Cooper, J. Foote, A. Girgensohn, and L. Wilcox, "Temporal event clustering for digital photo collections," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 1, pp. 269–288, Aug. 2005.

[81] C. Chen, M. Oakes, and J. Tait, "Browsing personal images using episodic memory (time+ location)," in *Advances in Information Retrieval*, pp. 362–372, Springer, 2006.

[82] K. Toyama, R. Logan, and A. Roseway, "Geographic location tags on digital images," in *Proceedings of the Eleventh ACM International Conference on Multimedia*, pp. 156–166, ACM, 2003.

[83] M. Naaman, S. Harada, Q. Wang, H. Garcia-Molina, and A. Paepcke, "Context data in geo-referenced digital photo collections," in *12th ACM conference on Multimedia*, pp. 196–203, ACM, 2004.

[84] A. R. Doherty, C. J. Moulin, and A. F. Smeaton, "Automatically assisting human memory: A sensecam browser," *Memory*, vol. 19, no. 7, pp. 785–795, 2011.

[85] A. R. Doherty, K. Pauly-Takacs, N. Caprani, C. Gurrin, C. J. Moulin, N. E. O'Connor, and A. F. Smeaton, "Experiences of aiding autobiographical memory using the sensecam," *Human–Computer Interaction*, vol. 27, no. 1-2, pp. 151–174, 2012.