# Experiences of Aiding Autobiographical Memory using the SenseCam

**Aiden R. Doherty[1,3], Katalin Pauly-Takacs[2], Niamh Caprani[1], Cathal Gurrin[1], Chris J.A. Moulin[2], Noel E. O'Connor[1], Alan F. Smeaton[1]**

[1] *CLARITY: Centre for Sensor Web Technologies, Dublin City University, Ireland*

[2] *Leeds Memory Group, Institute of Psychological Sciences, University of Leeds, U.K.*

[3] *British Heart Foundation Health Promotion Research Group, University of Oxford, U.K.*

## *Corresponding Author's Contact Information:*

aiden.doherty@dcu.ie

## ABSTRACT[1]

Human memory is a dynamic system which makes accessible certain memories of events based on a hierarchy of information, arguably driven by personal significance. Not all events are remembered, but those that are tend to be more psychologically relevant. In contrast, lifelogging is the process of automatically recording aspects of one's life in digital form without loss of information. In this article we share our experiences in designing computer-based solutions to assist people review their visual lifelogs and address this contrast. The technical basis for our work is automatically segmenting visual lifelogs into events, allowing event similarity and event importance to be computed, ideas which are motivated by cognitive science considerations of how human memory works and can be assisted. Our work has been based on visual lifelogs gathered by dozens of people, some of them with collections spanning multiple years. In this review article we summarize a series of studies that have led to the development of a browser which is based on human memory systems, and discuss the inherent tension in storing large amounts of data but making the most relevant material the most accessible.

---

1 Aiden Doherty is a computer scientist with an interest in behavioral lifestyle capture; he is a Marie Curie Postdoctoral research fellow in the Department of Public Health of the University of Oxford. Katalin Pauly-Takacs is a cognitive psychologist with an interest in human memory and its disorders; she is a final year PhD Research Student in the Institute of Psychological Sciences at the University of Leeds, UK. Niamh Caprani is a HCI researcher with an interest in technologies for older adults; she is an IRCSET PhD research student in CLARITY: Centre for Sensor Web Technologies in Dublin City University. Cathal Gurrin is a computer scientist with an interest in human digital memories; he is a lecturer at the School of Computing at Dublin City University in Ireland. Chris Moulin is a Cognitive Neuropsychologist interested in the organisation and awareness of human memory; he is a Senior Lecturer in the Institute of Psychological Sciences at the University of Leeds. Noel E. O'Connor is an electronic engineer with an interest in multi-modal sensor data analysis; he is an Associate Professor in the School of Electronic Engineering in Dublin City University and a Principal Investigator in CLARITY: Centre for Sensor Web Technologies. Alan Smeaton is a Professor of Computing and Principal Investigator in CLARITY: Centre for Sensoe Web Technologies. His research interests cover the management of all kinds of multimedia, and sensor, data.

**CONTENTS**

# 1. INTRODUCTION

*If we remembered everything, we should on most occasions be as ill off as if we remembered nothing.*
(James, 1890)

*Every piece of information is such that it is very unlikely, but just possible, that it is valuable.*
(O' Hara, Tuffield, & Shadbolt, 2008)

Autobiographical memory can be thought of as a store of the important events in our lives from which we construct our identity (for a review see Conway, 2005). As an example, people form collections of salient memories at times of identity formation, grouped around self images (Rathbone, Moulin, & Conway, 2008). For instance, our self image of being a parent will be supported by specific memories of important events such as birth, first steps, starting school and so on. The notion of a relationship between memory and the self is not new; Bartlett (1932) proposed that memory is not a mechanical process but a meaning-making system (see also Kant, 1798; Ribot, 1882).

Autobiographical memory has clear importance to daily life, personhood and well-being. Important events are preferentially retained in memory, for example relationships (McLean & Thorne, 2003) and events relevant to personal growth (Blagov & Singer, 2004). A critical psychological concept in lifelogging is nostalgia (for a review see Sedikides et al., 2008). Nostalgia is the willful accessing of autobiographical memories for positive outcomes; it enables continuity between the present self and one's personal past. A substantial empirical body of literature has shown that nostalgia generates positive affect, increases self-esteem, and fosters social connectedness (Wildschut et al., 2006). Designing tools to aid such activities should therefore be beneficial for the individual and society. Given the aging population, and therefore the increasing number of those likely to have a memory impairment (Van Den Broek, Cavallo, & Wehrmann, 2010), an important challenge for information scientists exists in developing technologies to aid autobiographical memory. Disruption to autobiographical memory has grave implications for personality (e.g. Addis & Tippet, 2004) and disorders which affect selfhood are accompanied by deficits in autobiographical remembering, such as depression (e.g. Dalgleish et al., 2007).

In this review summarizing over 5 years of work and publication, we consider how autobiographical memory may be supported through lifelogging, and we report our experiences with a device developed primarily for memory prosthesis, the SenseCam. Lifelogging refers to the digital capture of a person's everyday activities, in an unobtrusive and passive fashion. Apart from a few early visionaries (e.g. Bush, 1945) and pioneers (e.g. Mann, 1997) the field of lifelogging is a relatively new area of study. Much of the past research has focused on hardware miniaturization and storage (Mann, 1997; Aizawa, Ishijima, & Shiina, 2001). This has changed in the last 5-10 years with advances in storage, sensor and processor technologies leading to new digital recording and retrieval systems that may go beyond the views of the early visionaries (Bell & Gemmell, 2007). In the next 10 years a 250 terabyte hard drive (capable of holding tens of thousands of hours of video and tens of millions of photographs) may only cost $600, which should be enough to store all of the personal information encountered in an individual's lifetime. O'Hara et al. gives a good overview on what motivates us to investigate lifelogging activities:

> *"... Every piece of information is such that it is very unlikely, but just possible, that it is valuable. Before technology allowed comprehensive storage, our strategy was usually to try to estimate which information is likely to be more valuable and to keep that. Now there is no reason to stick to that philosophy…"* (O'Hara, et al., 2008).

The encoding specificity principle introduced by Tulving and Thomson (1973) states that information is best recalled when the cues present at capture match those that were present at encoding. In accordance with this, lifelogging devices that can be used to reinstate the visual context of personally experienced events may be best placed to support autobiographical memory. The most mature visual lifelogging device is the SenseCam, which was developed by Microsoft Research in Cambridge U.K., and is a wearable camera worn via a lanyard around the neck (see Figure 1). This device captures an image (approximately every 22 seconds) when triggered by sensors which log temperature, acceleration, light, and passive infrared data (Hodges, Berry & Wood 2011).

Unsurprisingly, the images captured using the SenseCam have been shown to operate as powerful autobiographical retrieval cues (Berry et al., 2009). However, much memory-focused lifelogging work has concentrated on those who are cognitively impaired, with positive results (e.g. Berry et al., 2009; Pauly-Takacs, Moulin, & Estlin, 2011). These studies focused on rehabilitation, and patients were instructed to wear the camera to record personally relevant or novel events; or patients were assessed completing experimental tasks, such as following a route. However our work differs in two key ways: 1) We consider tools to support memory in healthy people; and 2) we consider information access to very large, all-day every-day, lifelog collections gathered over extended periods of time.

****** FIGURE 1 ABOUT HERE ******

The SenseCam can capture up to 5,000 images and tens of thousands of sensor readings (e.g. accelerometer, lighting levels) in a busy day which can result in significant data volume. For example, one person who has worn the device on a daily basis has produced approximately 7,500,000 images in the last 5 years, each with the associated sensor information. As noted by others it is important that any tools facilitating navigation within this large collection of images should offer *"synergy not substitution"* of human memories (Sellen & Whittaker, 2010). This contrast between selective encoding in humans and constant recording in the SenseCam places a greater emphasis on retrieval processes and the organization of lifelogging materials to be presented to SenseCam users. In terms of human memory, information is often described as accessible or available (Tulving & Pearlstone, 1966). Available memory traces are those which have been successfully retained. Some information, for instance is processed at the time it occurs but is lost from memory, and is not available. Other information is available but not accessible – at least not until effectively cued or intervened by a retrieval strategy. As an example, the location of your friend's birthday celebrations last year can be momentarily inaccessible until cued by your friend reminding you that you had to wear an evening dress, at which point it becomes accessible. Extending such ideas to SenseCam, our operating principle is that *all* logged SenseCam images should be available, but echoing James above, not all should be accessible.

Here we summarize our work which has been driven by cognitive psychology principles and which have led to the construction of a platform to manage SenseCam images through exploiting or reflecting various characteristics of the human memory system. The remainder of this article is arranged around three major components of our work in providing lifelogging solutions to support human autobiographical memory:

1) **Event Segmentation:** the human mind reproduces memories in terms of events as the coherent units with a meaningful focus, thus SenseCam data should also be divided into events, where an event in this context refers to a specific activity of the wearer.

2) **Event Association:** as the human mind is largely driven by associative structures, so also should SenseCam events be easy to find and made accessible.

3) **Event Importance:** as distinct events are encoded more strongly in the human mind, we attempt to identify such distinct events and make them more accessible.

We will describe how each of these components has helped us evaluate the role of lifelogging in supporting personal recollection. Finally, we conclude with our experiences over the past five years, and consider the future challenges that lie ahead in developing software tools facilitating easy access to visual lifelog collections.

## 2. HUMAN MEMORY GUIDED COMPONENTS OF LIFELOGGING SOLUTIONS

### 2.1. Event Segmentation – Storing Images as Events

Human memory segments a continuously experienced present into a series of discrete events at retrieval (Williams, Conway, & Baddeley, 2008; Zacks, 2006). Despite the acceptance of this, there is very little experimental data to support theories of event segmentation. Williams et al. and Zacks both identify personal goals as the important drive towards segmenting experiences into events. For instance, looking back on any one day, one might identify an *'event'* as the journey to work. For this period, the primary goal is to arrive at work, and the collection of experiences and sensory information can all be related to this key goal. It is easy to determine the termination of the event according to the goal: once work has been reached, that goal is achieved, and the experiences are driven by a new goal and a new event is formed – for instance, making a cup of coffee. Arguably, events are hierarchical and somewhat driven by retrieval processes. Much more research is needed on event segmentation in human memory, and crucially, this is one area where SenseCam could be of great value where rich visual datasets available from ordinary daily lives can be used to determine the characteristics upon which event segmentation and identification is based.

## Event Segmentation Approach

Analogous to human memory, in lifelogging continuously experienced present (i.e. SenseCam images) should be segmented into a series of distinct events for later retrieval. An early problem encountered by the visual lifelogging community was in organizing and managing the millions of images produced by devices such as the SenseCam. Therefore an approach to managing lifelog images is to replicate how human memory works by merging clusters of similar images into discrete events (see Figure 2). This concept is not alien to the computing community as the traditional approach to content management for large video collections is to subdivide video (essentially a sequence of images) into '*shots*' (a grouping of similar images) (Smeaton et al., 2010).

******* Figure 2 about here *******

******* Figure 3 about here *******

The aim of automatic event detection from visual lifelogs is to determine boundaries that signify a transition between different activities of the wearer, whether visual, sensory, or otherwise. The journey to work, for instance, will create a unique signature of accelerometer and temperature data which will cease at the beginning of the next event – such as a more sedentary period at one's desk. The processes we formulated to achieve event segmentation can be summarized in four steps using only information from the SenseCam sensor data (without actually requiring any CPU intensive analysis of the images; see Figure 3):

**Compare various adjacent sensor (specifically motion) values against each other**

**to determine how dissimilar they are** – *higher degrees of dissimilarity indicate higher likelihood of a change in activity.* Firstly if one is talking to a friend but momentarily looks in the opposite direction an event boundary may be falsely triggered, therefore we compare aggregated 2-minute blocks of sensor values. As the sensor sources of information are all represented by single scalar values it is straightforward to compare sensor readings from adjacent readings. To calculate the difference between two accelerometer magnitude sensors readings, x and y, the answer is $D_l(x, y) = |x - y|$.

**Combine the various data sources together in an optimal manner** – *to verify there is agreement across the sensor sources that an activity change may be occurring.* After comparing adjacent image and sensor values against each other, there will be a separate list of difference values for each individual source. The greater the difference value, the greater the likelihood that an event boundary has taken place. Before data sources can be combined together it must be ensured that they are all on the same scale, using sum normalization (Montague & Aslam, 2001) where all values are shifted so that the minimum score is zero and the sum of all values summed together is one. Once the data sources have been normalized to a common scale, the process of combination, or fusion, can be carried out. We empirically determined that the CombMIM (Montague & Aslam, 2001) fusion approach is most suitable for lifelog event segmentation, where the minimum score from all the fused sources is taken, i.e. we only trigger an event boundary when the most doubtful source of information thinks an event transition has occurred.

**Determine a threshold value whereby higher dissimilarity values indicate areas that are likely to be event boundaries** – the magnitude of change required between activities must be sufficiently large to stop minor changes being suggested as events, but also not be too great where valid activity changes may not be registered. All the previous stages gave a likelihood of each instance being an event boundary between sequences of images, however no decision was made on which instances should be selected as the final event boundaries. We do this by automatically choosing a threshold value. If the threshold value is selected too low, there will be a number of false boundaries detected; however if the threshold is too high a value, there will be a number of valid boundaries undetected. We compared two thresholding techniques, one non-parametric (Kapur) and one parametric (Mean) (Sezgin & Sankur, 2004). We found the parametric mean-thresholding technique most suitable with an F1-measure of 0.6271 vs. 0.5799 for Kapur thresholding (Doherty, 2009). In Mean thresholding, the threshold is selected by adding k standard deviations to the mean, $T_{mean} = \mu + k\sigma$.

**Remove successive event boundaries that occur too close to each other** – *with the minimum event length empirically calculated to be 3 minutes*. At certain times some events may temporarily be interrupted by various distractions, where those distractions may not be long enough to merit being recorded as an autonomous event. Through experimentation we found that a gap of 3 minutes was best (Doherty, 2009). This also mirrors decisions made in public health behavioral understanding, where only episodes of greater than 3 minutes were used for active travel analysis (Kelly, 2011).

Early approaches to image clustering and segmentation either defined events as being of a fixed duration of time or were adaptations of approaches used to identify scenes in video (Wang, Hoffman, Cook, & Li, 2006; Yeung & Yeo, 1996) or stories from normal manually captured images. To evaluate the efficiency of our approach we compared it to four previous high-performing methods. These included the *Princeton Approach* (Wang et al, 2006) which segments lifelog videos into clips of fixed duration (5 minutes each); *Yeung & Yeo's* time constrained clustering technique; *RIAO* (Doherty et.al., 2007), an early prototype of the proposed event segmentation approach; and lastly we used the *sensors* within the SenseCam to define event boundaries (as just explained). Five participants were asked to collect free-living SenseCam images over a period of 1 month, with 61 days of valid data being subsequently collected. The participants then manually identified the boundaries between all events in their collection. This was achieved by having them look at all images for each day in sequence and then selecting where relevant transitions took place, using the SenseCam browser of (Hodges et. al., 2006). It was stressed to these users to judge event boundaries based on semantic meaning for that user personally. In total 2,986 boundaries were manually identified by the participants, giving an average of 19.1 events per day. We measured the performance based on precision, recall and a F1-Measure on 5 participants' SenseCam images with over 61 days of combined data. The current proposed event segmentation approach represented a 29.2% F1-measure improvement on prior work in the domain as illustrated in Figure 4.

******* Figure 4 about here *******

Our event segmentation process typically results in a full day's images (almost up to 5,000) yielding 20-30 events. The importance of the technique just presented means that we are using current state-of-art algorithms to present these events to end-users which can reduce the potential information overload in comparison to presenting all the images in an unclustered manner. Significantly this also structures SenseCam images (a continuously experienced present) into a series of discrete events which can then later be used at retrieval time.

## 2.2. Event Association – Associating Similar Events in Memory

Autobiographical memory relies on the integration of two stores, episodic memory and semantic memory. Episodic memory is event-based memories of specific instances, and is often characterized as mental *'time travel'* (it is also re-experienced with a sensation of *'remembering'*, indicating recollection). Semantic memory considers knowledge and conceptual representations (and does not give rise to the same subjective state). Conway and Pleydell-Pearce (2000) suggest that these two stores interact in a hierarchical manner in autobiographical memory. Autobiographical memories are essentially mental constructions incorporating event-specific episodic information into a factual, conceptual structure, termed semantic memory. For instance, we organize our memories into significant lifetime periods, relationships with others and even locations, within such a conceptual framework (e.g. *whilst I worked in Bath*) we can have specific memories (e.g. *the time I went to see Midsummer Night's Dream*). Since related events and specific instances are stored thematically it explains how when retrieving a specific memory from a particular lifetime period, other similar events tend come to mind, or are more accessible (see also Conway, 1996).

Based on this cognitive theory there is a clear need for lifelogging systems to provide users with automated search functions to find events similar to a given event, e.g. *"show me other times when I was at the park"*. Event association and retrieval in the domain of lifelogging has been investigated before, however experiments have been on very small datasets confined to just one user. At the most basic level a SenseCam event will consist of a number of images. Therefore to retrieve similar SenseCam events to a given event of interest in a lifelog it is necessary to firstly determine *how to represent* SenseCam events, and then *how to compare* those event representations against each other. We now discuss how this is achieved.

## Event Association Approach

Firstly an image can be described by its color, edge, and texture properties, in addition to a number of other traits. For our work we have used the standard MPEG-7 global color descriptors of *scalable color (64 element vector), edge histogram (80 element vector)*, and *color layout (12 element vector)* (Salembier & Sikora, 2002). On average each SenseCam event consists of almost 100 images. Each of those images is represented by a combined vector value, however it is desirable to obtain a single vector that is representative of the values of all 100 vectors. Smeaton & Browne (2006) note that in video retrieval the middle frame is often chosen to represent an entire shot (consisting of many images), this image is referred to as the keyframe image. However another means of representing an event is to combine multiple event images together into an average representative value, which can capture more of the elements of the event as a whole.

Having determined the method to represent each event visually, we then compare those event representative vectors against each other. The MPEG-7 features of each image (or event representation) are represented as a vector. After investigating 10 vector distance metrics (*Bray-Curtis, Canberra, Euclidean, Histogram Intersection, Jeffrey Modification of Kullback-Leiber, Kullback-Leiber, Manhattan, Square Chi Squared, Squared Chord, and $X^2$ Statistics*) we found it appropriate to use the standard Manhattan vector comparison method where $d_{man}(x, y) = \sum_{i=1}^{d} |x_i - y_i|$

To investigate the optimal approach to facilitate event search, we gathered another data collection of 273,744 SenseCam images from 4 participants (information retrieval specialists aged 25-35, wearing SenseCam for 1 month each). 50 events were selected as queries, and the users were then asked to judge a large number of potentially relevant events against each query event to build up a groundtruth of data. This was done in a TRECVid style pooling approach [154], with 43 possible system variations outputting their top 100 results for each query. Thereafter each user was presented with single keyframe images of all the unique pooled events, and asked to select those events they judged to be semantically relevant to the query image. This resulted in a user manually-defined groundtruth of 17,637 event-similarity-pairs. To have a sufficient number of relevant events to train parameters on, it was decided to go for more general queries in this dataset e.g. driving, at work on PC, eating, etc. This allowed us empirically investigate a number of event representation and event search scenarios. In an ideal world for event-event comparisons all images would have their visual features extracted, however this is computationally expensive and projecting forward towards the vision of ubiquitous lifelogging on the cloud, the scale of images produced (up to 5,000 per person per day) would merit an intelligent subset of images in each event. We empirically identified that by only extracting MPEG-7 features from the middle 35 images of each event (just over 30% of the entire set of images), the retrieval performance, in terms of MAP (mean average precision), is within 90% of when image features are extracted from all the images within an event. We recommend this approach be taken in future. We also found that while processing on sensor sources of information is very quick, image-based content information is needed for visual lifelog event search purposes. A disadvantage of the event-similarity-pairs dataset is that while it was necessary to select very general queries to produce a sufficient number of relevant events on which to tune retrieval parameters, these queries are not representative of all possible user query classes.

Therefore we decided to create a second dataset on which users were asked to construct real world queries with very specific information needs. In experiments to investigate the effectiveness of our retrieval approaches for real user generated queries on extensive datasets, we asked four users to collect SenseCam data over a period of at least one month. A total of 1,864,149 SenseCam images were used in this experiment, which were automatically segmented into 22,125 events (using approach described in Section 2.1). The users identified 23 query events from which they'd like to find other similar events to. We then used the CPU intensive approach of using all images in an event to compare event-representative vectors. Unfortunately while retrieval performance was encouraging in the event-similarity-pairs dataset of generic queries (% of top 5 ranked results which are relevant, P@5 = 0.69), the performance on the 23 specific queries on the larger dataset is insufficient (P@5 = 0.30). To verify that those results were using state-of-art multimedia retrieval techniques, we compared our global color based-approach just mentioned to methods for extracting interest point features from images using SIFT [Scale Invariant Feature Transform, (Lowe, 2004)] and SURF [Speeded Up Robust Features, (Bay et al., 2006)] techniques. Figure 5 illustrates that our algorithm is at least as comparable, thus meaning that a significant challenge remains for the community to improve the performance of specific user queries of interest.

Previously the dominant presentation paradigm for reviewing lifelog images was a conventional sequential "replay" or fast-forwarding of all captured images, more formally known as RSVP (Spence, 2002). Other early visualization methods included integration with GPS devices (Bell, Gemmell, & Gates, 2010). Through the process of event segmentation, and event comparison we introduced visual search functionality for end-users to browse through their lifelog collections (Lee et al., 2008). This functionality can support the association of events with each other. In theory this computational semantic association may complement human episodic memory. For example if images exist of playing in the snow, other similar existing events can automatically be associated with this, using the event representative MPEG-7 vectors and distance comparison techniques just mentioned. However given the current poor performance of the state-of-art in automatically identifying specific queries of interest (P@5=0.30), there remains two challenges to the community: 1) Improving automatic event association performance; 2) Developing user-conscious semi-automatic query functionality.

******* Figure 5 about here *******

## 2.3. Event Importance – Considering Self Issues

An explanation for the poor performance on specific queries of interest is that there is simply a smaller pool of relevant events that can be found, rather than the greater likelihood of retrieving more routine events. Another reason is due to the computational challenge which exists in automatically assigning semantic meaning or significance to images which are essentially stored as color pixels. This is the long-identified '*semantic gap*' where computers fail to translate bytes into real semantic meaning (Smeulders et al., 2000). Attempting to determine the *'meaning'* or *'significance'* of an event is a subjective exercise. In this case, our attempt to consider meaning in memory is to examine what gets preferentially stored, and therefore what is relevant to the self. In doing so, we focused on two factors: face detection (as a proxy for social importance) and distinctiveness.

In the nostalgia literature, social events and significant relationships feature predominantly as personally relevant autobiographical memories (Wildschut et al., 2006). Early efforts in the lifelogging domain followed this principle where social interaction and conversational scenes were regarded as key elements in determining event importance. To detect these, automatic face detection was used to determine events containing face-to-face conversations. Within each event all images are firstly processed to investigate whether a face is present or not using the Intel OpenCV face detection toolkit (*haarcascade-frontalface-alt, scaling factor of 1.1, 3 neighbors, and window size of 30 pixels*). Thereafter each event is given a face concentration score, with the count of all images with >=1 faces present being divided by the total number of images in the event.

In recognition memory tasks using experimentally controlled sets of words or pictures, participants are better able to detect distinctive items (e.g. Mandler, 1980). Distinctiveness is often encountered with reference to novelty, in that novel events are more distinctive. It is widely assumed that novelty is a major trigger for orientating attention to a stimulus in the environment, and that this novelty engages memory encoding mechanisms (e.g. Grunwald & Kurthen, 2006). Distinctiveness is also a critical issue in autobiographical memory (Brewer, 1988). Research has considered which events from across the lifespan are better remembered. If we plot the accessibility of memories across the lifespan, a robust phenomenon called the reminiscence bump is produced. This is a period in life, typically between the ages of 15 and 25, where a great amount of memories are produced (e.g. Rubin, Wetzler, & Nebes, 1986). There are various accounts of this phenomenon, but one suggestion is that it is driven by distinctive, important, self-defining first-time experiences; for instance, embarking on a career, meeting a life partner, achieving academic qualifications, and so on. Extreme cases of distinctiveness concern *'flashbulb memories'* (e.g. Brown & Kulik, 1977). These are events which are sufficiently distinctive so as to produce extremely vivid memory of the circumstances of an event. Typically, these are memories of public events, such as the attack on the World Trade Center of 2001. In such instances, it appears that distinctiveness serves to make the memory of hearing this news very memorable, even after long periods and in neurodegenerative conditions. Arguably, these events are of personal and public significance and this drives the superior memory for them.

Considering that more distinctive events are better remembered in human memory, it follows that a browser should be able to present events to the user that are more interesting on the basis of their distinctiveness. For human memory to be supported by lifelogging systems, those events that are likely to be better remembered by an individual should be automatically identified. As a case of further motivation, consider that a user will on average capture over 7,000 events per year (assuming approximately the lower-case scenario of 20 events per day), automatically summarizing the collection to the more interesting events will support the user reflect upon their experiences.

## Event Importance Approach

Given we have just motivated that distinct events are more strongly remembered, in 2008 we extended previous lifelogging efforts by introducing the notion of event novelty whereby visually distinct or outlier events are likely to be more distinctly and strongly remembered. Firstly MPEG-7 image descriptors are extracted from the middle 35 images from each event (same as used in Section 2.2), and thereafter the event's novelty score is the sum of its Manhattan distance metric against all relevant comparison events, divided by the count of those comparison events. Less frequently occurring (i.e. more novel) events will have higher distance metric scores in comparison to routinely occurring ones which will have lower similarity scores. We empirically found the most appropriate set of events on which to compare a given SenseCam event to, were all events that occurred +-2 hours on the same day of the week in the entire history of the dataset. The premise for this is to highlight, for instance, a family meal out in a nice restaurant on a Thursday evening rather than the normal evening meal at home.

******* Figure 6 about here *******

Due to lifelogging technologies only recently becoming more easily available, three years ago we were not in a position to carry out life-duration experiments, so we firstly evaluated the effectiveness of our event importance approach on ranking the importance of events within a day. Three information retrieval specialist users (aged 20-35) wore the SenseCam for one month each (total of 176,975 images segmented into 1,758 events). As there is a subjective nature of rating how important an event is in relation to other events, it is very difficult to rank the importance of all events within a day, and to do this for each and every day would present a large annotation burden on users. Therefore given that it is of much interest to determine the most interesting events in a given day, in addition to determining the most mundane/routine events from a day, a decision was made to present keyframe images of all the day's events, and then the two most important and two least important (as determined by the approach under investigation) events to the user (Doherty, 2009, pp. 172). Users were then asked to give a single Likert judgment on how much they agree with the proposed most and least important events as a summarization of that day. Future investigations in this area could look at analyzing the most and least important events separately, and then the interaction between them. However this analysis was not the aim of our experimental design at the time. After 664 judgments made by our users, we found the most effective approach is in combining the automated detection of faces (to indicate social engagement) with detecting how visually novel each event is. To obtain the novelty score, each event in a day is automatically visually compared (via distance metrics between event representative MPEG-7 vectors) to see how dissimilar it is to other events occurring +/- 2 hours in previous same weekdays. Figure 6 displays the success (in terms of typical Likert ratings on a scale of 1-5) each of the systems had in identifying events of interest that closely matched those defined by each user.

In summary we have identified that the human mind groups together continuous material into discrete events, that those events have various degrees of association with past events, and the most distinct events are easiest to retrieve. As described, computational processes have been developed to mimic those functions of human memory. The meaning of this is that lifelogging solutions understanding how the human mind is likely to operate, should in theory be well placed to support autobiographical functioning and personal recollection.

## 3. PERSONAL RECOLLECTION – HOW APPROPRIATELY DESIGNED LIFELOGGING TOOLS CAN ASSIST

Brewer (1988) was the first to scientifically investigate ‘personal recollection’, where diary methods were used to try and gauge what people could retrieve from daily life over very short sections of the lifespan. Visual lifelogging has since opened up new possibilities to carry out more ambitious studies in this field, where it is now possible to carry out personal recollection experiments on very large lifelog collections. For example Sellen et al. (2007), has shed light on how SenseCam interacts with different forms of human memory. In their study, participants were asked to classify their subjective experience on recalled events from SenseCam days (when they wore SenseCam) and control days (when they did not wear SenseCam) as *remembered* (reflecting episodic memory) or simply *known* (reflecting semantic memory). The key issue in studies of this type is that participants can reliably differentiate memories which are ‘remembered’ from which are ‘known’ (e.g. Conway, 2005). Firstly, it was found that reviewing SenseCam images gave rise to higher recall in both types of memories than did control images. Recall in this study was measured over two minutes by asking the participants to write prose cued by the questions ‘what’, ‘where’, ‘when’, and ‘who’ for each event. Furthermore, whereas SenseCam-cued *remembered* events gradually became more and more difficult to recall over time, SenseCam-cued *known* events showed greater stability. This suggests that if details of a SenseCam-recorded event are no longer available for conscious recollection due to general forgetting effects, the knowledge of having experienced that event will still be available. Contrastingly, *known* events of control days (no SenseCam) did significantly decline over time, which confirms the role of the device in the long-term retention of *knowing* that an event did occur in one's past. However, in this study participants only wore the SenseCam for a very short time-period of two weeks. In this article we consider the scenario of interacting with a lifelog collection of years in duration, and how software tools can assist personal recollection in such a scenario.

Inspired by studies of recollection in the laboratory which ask participants to report their experience according to a number of basic questions (who, where, what; Perfect et al., 1996) we designed a new lifelogging browser to acknowledge that there are multiple sensory routes on which events can be associated. This led to the construction of a system, illustrated in Figure 7, where the user could query by the following search axes: *where* (location, altitude, temperature); *when* (calendar selection, prev/next day browsing, season, year, day/night, time of day, & month); *what* (visual appearance, bright/dark, important/routine, semantic concepts (Doherty et al., 2011) e.g. eating, working on PC, etc.); and *who* (estimated number of people in scene based on face detection).We now briefly share our thoughts on what event association strategies may best support personal recollection based on a single case, a permanent SenseCam wearer 34-year-old male (Moulin et al., in prep). Given the prior lack of availability of lifelogging devices it has only been possible to gather a multiyear data collection from one committed individual willing to share his images (author CG). Ideally, a full group study would be conducted, but we had a unique opportunity to research his then 2.5 million (now 7.5 million) image multi-year lifelog data collection, in a single case design. Our view was that CG presented to us a 'natural experiment' in the same way that special cases of brain injury We feel that the lessons learned from the experiences working with this one user can help inform the community on future directions for when lifelogging becomes mainstream. Our volunteer participant - CG, a computer scientist by profession - had gathered a collection of 2,579,455 SenseCam images over a time period of 2.5 years. Our event segmentation (Section 2.1) divided these images into 29,301 events. Event importance technology (Section 2.3) then allowed the selection of potentially interesting events from this collection of distinct events. Using the lifelog browsing system displayed in Figure 7, our case study concentrated on 1) how successfully the system could identify events of personal importance to the user; and 2) how the system could support the user to find events of interest. CG generated his 50 specific events from the 2.5 year time period that he had been wearing SenseCam. These 50 events were obtained employing a version of an autobiographical fluency task (e.g. Dritschel et al., 1992), whereby CG freely generated as many memories as he could as quickly as possible, no instruction was given about time period, topic, or personal significance. He gave titles and dated each memory as it came to mind, and then moved onto retrieving the next. Once all 50 events were retrieved, he then rated the novelty and personal importance of events on a scale from 1 to 7. CG was able to retrieve the 50 events without much difficulty such that the task took approximately half an hour to complete.

CG's set of memories allowed us to investigate the success of the lifelog system's event importance module in identifying the 50 most interesting events from the participant's 29,301 events. To enable this, CG also rated the novelty and personal importance of the lifelog system's 50 most important events, plus an additional 50 randomly selected control events. The mean ratings for these events were submitted to analysis of variance. There was an effect of event type (CG-generated events versus lifelog's important events versus random events) for both novelty and personal importance, $F(2, 151) = 38.8$, $p<.001$ and $F(2, 151) = 20.5$, $p<.001$., respectively, such that each event type was significantly different from each other (all at $p < .01$). More specifically, CG rated his self-generated memories as being the most novel and personally significant events, but the lifelog system's events were rated significantly more novel and personally important than the random (control) generated events.

******* Figure 7 about here *******

Typically, CG's highest ratings were given to family events (e.g. wedding), to events pertaining relationships (e.g. meeting girlfriend), and to events signifying a change in lifestyle (e.g. buying a car or a new home). The outcome is that the lifelog system generated items which were significantly more novel and personally significant than a control set of random events, but which did not choose events as novel or as personal significant (as rated by the user). Such a method could be taken forward as a way of examining the relationship between computer-generated important events and those which are subjectively the most important.

Despite some success in making automatically accessible novel and significant events, perhaps it is expecting too much of a lifelog system to know which events to make the most accessible with complete overlap. A more modest aim would be to have a lifelog system where the subjectively important and novel events were at least easily accessible. To this end, approximately 6 months after our initial testing on CG, we asked him to find his 50 self-identified events in the system, and recorded his personal reflection of those events. Two systems were designed to help CG retrieve these 50 events, the first being retrieved using our old lifelog browser, which only offers time (via calendar) and visual similarity (via side panel) based search, (Doherty, Moulin, & Smeaton, 2011), and the second being retrieved using the lifelog browser in Figure 7; The browsers cannot be compared on the same set of queries since learning effects would influence the results. The data suggests that our lifelog browser aided our participant to retrieve his 50 most interesting events with the median search time of 127 seconds (38 events retrieved) on the collection of 29,301 events, as opposed to 774 seconds (12 events retrieved) using the older lifelog browser. After just 12 queries it was clearly obvious that the old browser was ineffective, so we therefore searched the remaining queries using the lifelog browser described in this article. Offering people multiple sensory paths on which to access their lifelog collection suggests early promise, but merits further investigation.

## 4. CONCLUSIONS AND REFLECTIONS – WHERE NEXT

We have reflected on various instances of how biomimicry of the human autobiographical memory system has resulted in significant gains achieved in lifelogging systems. Over the past five years we have carried out dozens of experiments on approximately 15 million SenseCam images captured by over 40 different participants. We now conclude by reflecting on our past experience, and look forward to future research directions the lifelogging community should take to support healthy individuals in reviewing captured visual records.

### Reflecting on the Past Five Years

A perception has existed that the field of lifelogging has been overly focused on recording the minutiae of everyday life, but without making the data meaningful to end users (Sellen & Whittaker, 2010). However our experience of working with real-world users has been that by capturing as much data as possible, we can better direct the wearer towards significant moments in their lives, i.e. the redundancy of everyday mundane events assists in identifying the outlier and more memorable events. In essence, capturing everything does not mean that we must review everything, indeed forgetting is very important, but capturing as much as possible creates the best environment to guide us towards those everyday moments that are significant in our lives. We believe that the tools developed from our experiences (e.g. http://sensecambrowser.codeplex.com) will support the personal memories community design better reflection and interaction methods (e.g. Petrelli et al., 2009).

The sensor sources of information are valuable to assist the computational processing of event segmentation, but for event search the images are still most powerful to induce the recollective experience (also recognized by Kalnikaitė et al., 2010). We have also learned that it is not necessary to strictly define boundaries between events, as the events act as a quick navigation towards the images of interest. This exploits the human capabilities of gisting or inferring what an event was about, and also inferring the temporal ordering of autobiographical events (Brainerd & Reyna, 2001; Koriat, Goldsmith, & Panksy, 2000)

On an individual's lifelog of 2.5 years, anecdotal evidence indicates that the average time to find an event (among the nearly 30,000 present) has been reduced from 774 seconds when browsing to 127 seconds when presenting the user with multiple sensory paths. However 127 seconds to find a target event is not acceptable to users who would naturally expect prompt access to relevant information from their lifelog. This still represents the single greatest challenge for our community.

## Looking Forward to the Future

The commercial release of the SenseCam, via the branding of the Vicon Revue, is important in creating availability of equipment, which will enable studies to be carried out on larger and more diverse populations. Hardware no longer poses a significant issue, and neither do storage and processing (Doherty, Moulin, & Smeaton, 2011). The next computational/technology challenge lies in semantic interpretation and search. This is a process which requires the guidance of psychological principles and an understanding of what motivates self-driven goals. This fundamental search work will provide the platform for the next generation of digital personal memory reflection tools, just as the past five years have driven a suite of studies.

In the video retrieval domain benchmarking exercises have been instrumental in extending the state-of-art performance (Thornley et al., 2011). However given the early stage of research in the lifelogging domain, and some initial concern surrounding the sharing of participants' automatically captured lifelog images of non-consenting individuals (Allen, 2007), there has been no benchmarking dataset made available yet. For the search performance to significantly improve, a common dataset on which to carry out benchmarking exercises will be essential, notwithstanding the challenges of generating a suitable dataset given the personal nature of the data.

Given the role that lifelogging could have on society, the lifelogging community should involve *'reflectors'* from the arts and humanities throughout the process of technological advancement. Such stakeholders could play an integral role in ensuring that the potential technological benefits of lifelogging sufficiently outweigh the perceived sociological implications (Allen, 2007). The anticipated benefits include scenarios of social sharing of experienced events or happenings as recorded using lifelogging technologies, supporting human memory, preserving the experiences of a loved one long after their passing away, and many more. Given that potential implications of lifelogging technologies could include breaches of privacy of the individual, issues of control of content, regulation governing ownership of lifelog data after death, laws regarding forced sharing of lifelog data to resolve legal disputes, etc. These concerns should be addressed.

Although the benefit of lifelogging technologies such as SenseCam are emphasised in people with memory difficulties, we feel users (particularly older users) would be more likely to adopt lifelogging technologies to support their memory if they have previous experience using them prior to their impairment. Therefore it may be necessary to design a lifelogging application that older adults would be motivated to use. Some motivational factors could include: integration of information that older adults are already interested in; emphasising family collaboration in lifelogging; and supporting storytelling and reminiscence.

Finally, lifelogging issues a challenge to the memory community. Given that the nostalgia literature (e.g. Wildschut, et al., 2006) shows that there is a benefit to wellbeing and mood by freely retrieving memories from the past, what will the effect be when memory is supported through lifelogging? This can only be answered with a longitudinal study where enough of one's lifespan has elapsed in the lifelog to emulate the long-term recollections of important events seen in the reminiscence bump.

## In Summary

We began this paper by outlining two quotes which sum up what we feel are the central issues for the lifelogging community, the psychological redundancy of remembering everything contrasting with the technological ability to store everything. Even though O'Hara et al.'s (2008) comment is quite recent, the sentiment has long been a driving force of lifelogging efforts. However only five years ago the comments of James (1890) were viewed as most practical by the wider community. Through a process of biomimicry of the human memory system and then developing technologies to complement autobiographical memory we feel that the wider view is now shifting towards O'Hara et al.'s comments. It is our aspiration that future lifelogging efforts set up rigorous experiments to conclusively answer this debate, somewhat analogous to the effect the TRECVid benchmarking exercise has had on video search advancements (Thornley et al., 2011). To refer back to the literature, we are in a position where all material in lifelogs should be available, but our browsers and retrieval systems should consider how to make accessible information which is easy to search, relevant, and perhaps most critically, be of personal significance.

## NOTES

***Acknowledgments.*** This material is based upon works supported by the Science Foundation Ireland under Grant No. 07/CE/I1147 and the Irish Research Council for Science Engineering and Technology. The authors also wish to acknowledge the support of Microsoft Research (grant number 2007-056). Author AD is a Marie Curie postdoctoral research fellow supported by the Irish Health Research Board under grant number MCPD/2010/12.

***Support.*** An open source version of the SenseCam browser, featuring event segmentation described here is available at: http://sensecambrowser.codeplex.com/

***Authors' Present Addresses.*** Aiden R. Doherty, CLARITY: Centre for Sensor Web Technologies, Dublin City University, Ireland. Email: aiden.doherty@dcu.ie

***HCI Editorial Record.*** (supplied by Editor)

## REFERENCES

Addis, D. R., & Tippett, L. J. (2004). Memory of myself: Autobiographical memory and identity in Alzheimer's disease. *Memory*, *12*, 56–74.

Aizawa, K., Ishijima, K., & Shiina, M. (2001). Summarizing wearable video. Proc. International Conference on Image Processing, 3. 398-401.

Allen, A.L. (2007). Dredging-up the past: Lifelogging, memory and surveillance. *New Yorker*, 38-44.

Bartlett, F. C. (1932). Remembering. A study in experimental and social psychology. Cambridge University Press

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF:speeded up robust features. *In Proceedings of the 9th European Conference on Computer Vision (ECCV'06)*.

Bell, G., & Gemmell, J. (2007). A digital life. *Scientific American*, 296(3), 58-65.

Bell, G., Gemmell, J. (2010). Your life, uploaded: The digital way to better memory, health, and productivity. New York, NY: Penguin Group.

Berry, E., Hampshire, A., Rowe, J., Hodges, S., Kapur, N.,Watson, P., et al. (2009). The neural basis of effective memory therapy in a patient with limbic encephalitis. *Neurology, Neurosurgery, and Psychiatry with Practical Neurology*, 80 (3), 582–601.

Blagov, P.S., & Singer, J. A. (2004). Four dimensions of self-defining memories (specificity, meaning, content, and affect) and their relationships to self-restraint, distress and repressive defensiveness. *Journal of Personality, 72* (3) pp. 481–511.

Brainerd, C.J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behaviour*, 28, 41–100.

Brewer, W.F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21-90). Cambridge: Cambridge University Press.

Brown, R. & Kulik, J. (1977). "Flashbulb memories", *Cognition, 5* (1), 73-99.

Bush, V. (1945). As we may think. *The Atlantic Monthly, 176*(1), 101–108.

Conway, M.A. (1996). Autobiographical knowledge and autobiographical memories. In D.C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67-93). Cambridge: Cambridge University Press.

Conway, M.A. (2005). Memory and the self. *Memory and Language, 53* (4), 594 – 628.

Conway, M.A., & Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107*(2), 261–288.

Dalgleish, T., Williams, J.M.G., Golden, A.J., Perkins, N., Barnard, P.J., Au-Yeung, C., Watkins E. (2007). Reduced Specificity of autobiographical memory and depression: The role of executive processes. *Journal of Experimental Psychology: General*, *136 (1)*, 23-42.

Doherty, A.R., Smeaton, A.F., Lee, K., & and Ellis, D.P.W. (2007). Multimodal segmentation of lifelog data. *Proceedings RIAO 2007 - Large-Scale Semantic Access to Content (Text, Image, Video and Sound)*, Pittsburg, PA, USA, 2007.

Doherty, A.R. (2009). Providing effective memory retrieval cues through automatic structuring and augmentation of a lifelog of images. PhD thesis, Dublin City University.

Doherty, A.R., Moulin, C., & Smeaton, A.F. (2011). Automatically Assisting Human Memory: A SenseCam Browser *Memory: Special Issue on SenseCam, The Future of Everyday Research?, 19(7),* 785-795.

Doherty, A.R., Caprani, N., Ó Conaire, C., Kalnikaite, V., Gurrin, C., Smeaton, A.F. & O'Connor, N.E. (2011). Passively recognising human activities through lifelogging. *Computers in Human Behavior*, *27(5)*, 1948-1958

Dritschel, B. H., Williams, J.M.G., Baddeley, A. D., & Nimmo-Smith, I. (1992). Autobiographical fluency: A method for the study of personal memory. *Memory & Cognition*, *20 (2)*, 133-140.

Grunwald T. & Kurthen M. (2006). Novelty detection and encoding for declarative memory within the human hippocampus. *Clinical EEG and Neuroscience 37(4)*, 309–314.

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., ...Wood., K (2006). SenseCam: a retrospective memory aid. In UbiComp: 8th International Conference on Ubiquitous Computing, vol. 4602 LNCS, pp 177–193. Springer.

James, W. (1890). *The principles of psychology*. New York: Henry Holt.

Kalnikaitė, V., Sellen, A., Whittaker, S., & Kirk, D. (2010). Now let me see where I was: Understanding how lifelogs mediate memory. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2010) Atlanta, Georgia, USA, April 13-19, 2010.

Kant, I. (1798). Anthropology from a pragmatic point of view. (*Anthropologie in pragmatischer Hinsicht*)

Koriat, A., Goldsmith, M., & Pansky, A. (2000).Toward a psychology of memory accuracy. *Annual Review of Psychology*, 51, 481–538.

Kelly, P., Doherty, A.R., Berry, E., Hodges, S., Batterham, A., Foster, C. (2011). Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. *International Journal of Behavioral Nutrition and Physical Activity 2011, 8:44*

Lee, H., Smeaton, A.F., O'Connor, N.E., Jones, G.J.F., Blighe, M., Byrne, D., Doherty, A.R. & Gurrin, C. (2008). Constructing a SenseCam visual diary as a media process. *Multimedia Systems Journal*, *14* (6), 341-349. ISSN 1432-1882

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60* (2), 91-110.

McLean, K. C., & Thorne, A. (2003). Adolescents' self-defining memories about relationships. *Developmental Psychology, 39,* 635–645.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review 87*, 252-71.

Mann, S. (1997). Wearable computing: a first step toward personal imaging. *Computer*, 30, 25–32.

Montague, M., & Aslam, J. A. (2001) Relevance score normalization for metasearch. In CIKM '01: Proceedings of the tenth international conference on Information and knowledge management, pages 427–433, Atlanta, Georgia, USA, 2001.

Moulin, C.J.A., Pauly-Takacs, K., Doherty, A., Gurrin, C., Smeaton, A. (in preparation). Three years of Lifelogging data from SenseCam: Insights into autographical memory.

O'Hara, K., Tuffield, M. & Shadbolt, N. (2008, May). Lifelogging: Issues of identity and privacy with memories for life. Paper presented at the First International Workshop on Identity in the Information Society, Arona, Italy. Retrieved from http://eprints.ecs.soton.ac.uk/15993/

Pauly-Takacs K., Moulin, C. J. A. & Estlin, E. J. (2011). SenseCam as a rehabilitation tool in a child with anterograde amnesia. *Memory*, *19*, 705-712.

Perfect, T. J., Mayes, A. R., Downes, J. J., Van Eijk, R. (1996). Does Context Discriminate Recollection from Familiarity in Recognition Memory? *The quarterly journal of experimental psychology, 49A (3)*, 797-813.

Petrelli, D., van den Hoven, E., Whittaker, S. (2009). Making history: Intentional capturing of future memories. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2009) Boston, USA, April 4-9, 2009, pp.1723-1732.

Rathbone, C.J., Moulin, C.J.A., Conway, M.A. (2008). Self-centered memories: The reminiscence bump and the self. *Memory and Cognition, 36(8),* 1403-1414.

Ribot, R. (1882). *Diseases of Memory*. New York: Appleton

Rubin, D. C., Wetzler, S. E., & Nebes, R. D. (1986). Autobiographical memory across the lifespan. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 202-221). Cambridge: Cambridge University Press.

Salembier, P., & Sikora, T. (2002) *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley & Sons, Inc., New York, NY, USA.

Sedikides, C, Wildschut, T, Arndt, J. & Routledge, C, (2008). Nostalgia: past, present, and future. *Current Directions in Psychological Science, 17 (5),* 304-307.

Sellen, A. & Whittaker, S. (2010). Beyond total capture: A constructive critique of lifelogging. *Communications of the ACM, 53*, 70-77.

Sellen, A.J., Fogg, A., Aitken, M., Hodges, S., Rother, C., & Wood, K. (2007). Do life-logging technologies support memory for the past?: An experimental study using sensecam. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2007) New York, USA, April 28- May 3, 2007, pp. 81-90.

Sezgin, M. & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electonic Imaging, 13(1),*146–168.

Smeaton, A.F. & Browne, P. (2006). A usage study of retrieval modalities for video shot retrieval. *Information Processing and Management, 42*(5), 1330-1344.

Smeaton, A.F., Over, P, & Doherty, A.R. (2010). Video shot boundary detection: seven years of TRECVid activity. *Computer Vision and Image Understanding, 114 (4)*. pp. 411-418. ISSN 1077-3142

Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22,* 1349-1380.

Spence, R. (2002). Rapid, Serial and visual: a presentation technique with potential. *Information Visualization 1(1)*, 13-19.

Thornley, C., Johnson, A.C., Smeaton, A.F., & Lee, H. (2011). The scholarly impact of

TRECVid (2003-2009). *Journal of the American Society for Information Science,* 62*(4),* 613-627.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, 25, 1-12.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Verbal Learning and Verbal Behavior*, 5, 381 – 391.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80(5),* 352–373.

Van Den Broek, G., Cavallo, F., & Wehrmann, C. (2010). *AALIANCE ambient assisted living roadmap (vol. 6).* Amsterdam, Netherlands: IOS Press.

Voorhees, E. M. (2007). TREC: Continuing information retrieval's tradition of experimentation. *Communications of the ACM, 50(11),* 51–54.

Wang, Z., Hoffman, M.D., Cook, P.R., Li, K. (2006). VFerret: Content-based similarity search tool for continuous archived video. *Proceedings, CARPE '06: ACM Workshop on Continuous Archival and Retrieval of Personal Experiences,* pp 19-26, October 2006.

Wildschut, T., Sedikides, C., Arndt, J., & Routledge, C. D. (2006). Nostalgia: Content, triggers, functions. *Journal of Personality and Social Psychology,* 91, 975-993.

Williams, H.L., Conway, M.A., & Baddeley, A.D. (2008). The boundaries of episodic memories. In T.F. Shipley & J. M. Zacks (Eds), Understanding events from perception to action. (pp 589-616). New York: Oxford University Press.

Yeung, M.M. & Yeo, B.L. (1996). Time-constrained clustering for segmentation of video into story units. *Proceedings, International Conference on Pattern. Recognition, Vol. C*, pp. 375-380, August 1996.

Zacks, J.M., Speer, N.K., Vettel, J.M., & Jacoby, L.L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging*, 21, 466–482.

## FIGURE CAPTIONS

**Figure 1. The Microsoft SenseCam digital camera**

**Figure 2. Segmenting SenseCam images is analogous to how human memories stores continuously experienced material as discrete events**

**Figure 3. Overview of process to identify transitions between lifestyle events from SenseCam data, using the sensor sources only**

**Figure 4. Performance in identifying boundaries between lifelog events. The x-axis sorted in descending order of performance of our "Sensors Only" approach represents 61 discrete days from 5 individuals, and the y-axis represents the event-segmentation F1-measure accuracy obtained for each of those 61 days. Our method in the thick black line represents an improvement over existing techniques**

**Figure 5. A summary of the MAP (mean average precision) performance of our approach (MPEG7Sense) in comparison to using elementary SIFT and SURF interest feature comparisons on 23 queries of specific interest to 4 users**

**Figure 6. Identifying events that are likely to be more distinctly encoded in the autobiographical memory system**

**Figure 7. Visual lifelogging "multi-axes" browser developed in 2010. The primary design goal was to provide multi-faceted retrieval of events to support person recollection. This browser aims to support search on the "who", "what", "when" and "where" axes of retrieval**

## FIGURES

### Figure 1. The Microsoft SenseCam digital camera

**Figure 2. Segmenting SenseCam images is analogous to how human memories stores continuously experienced material as discrete events**

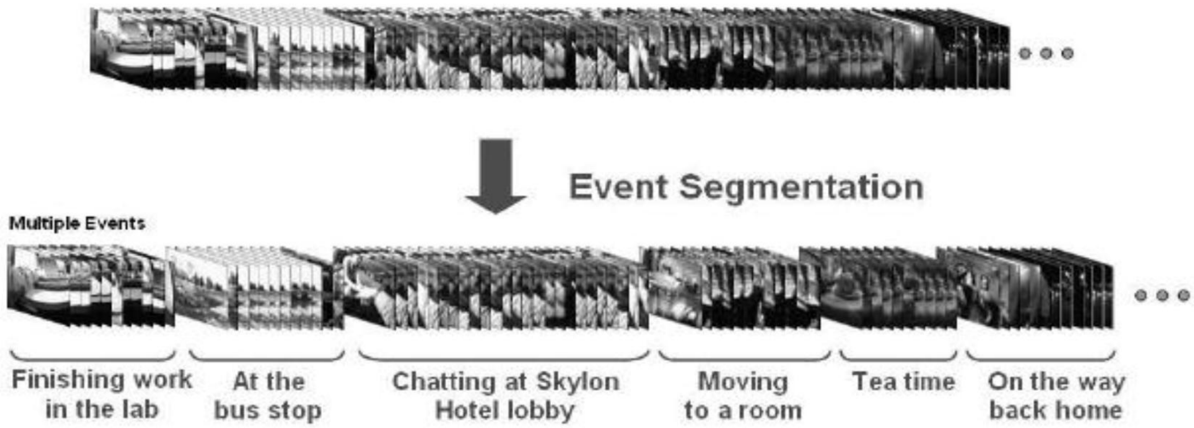

**Figure 3. Overview of process to identify transitions between lifestyle events from SenseCam data, using the sensor sources only**

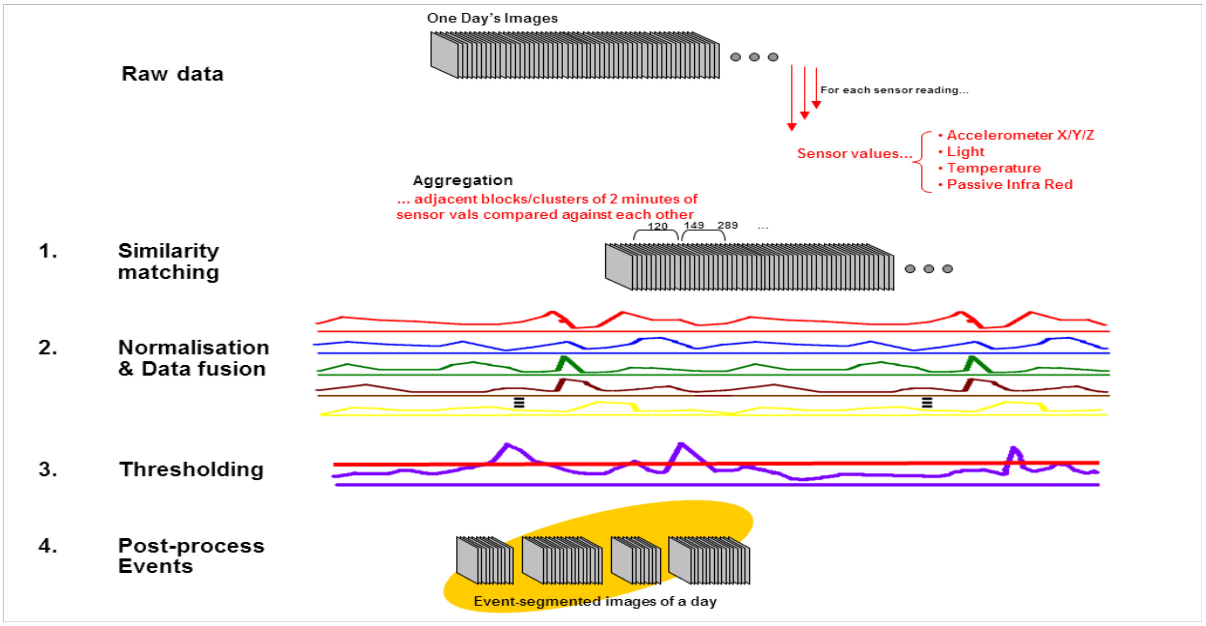

**Figure 4. Performance in identifying boundaries between lifelog events. The x-axis sorted in descending order of performance of our "Sensors Only" approach represents 61 discrete days from 5 individuals, and the y-axis represents the event-segmentation F1-measure accuracy obtained for each of those 61 days. Our method in the thick black line represents an improvement over existing techniques**

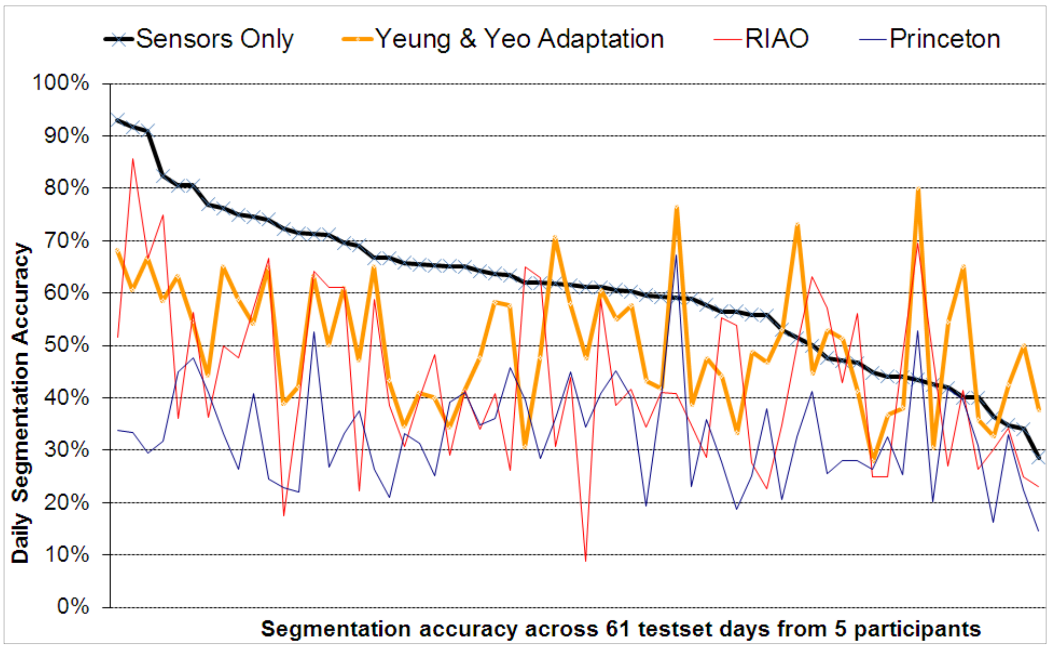

**Figure 5. A summary of the MAP (mean average precision) performance of our approach (MPEG7Sense) in comparison to using elementary SIFT and SURF interest feature comparisons on 23 queries of specific interest to 4 users**

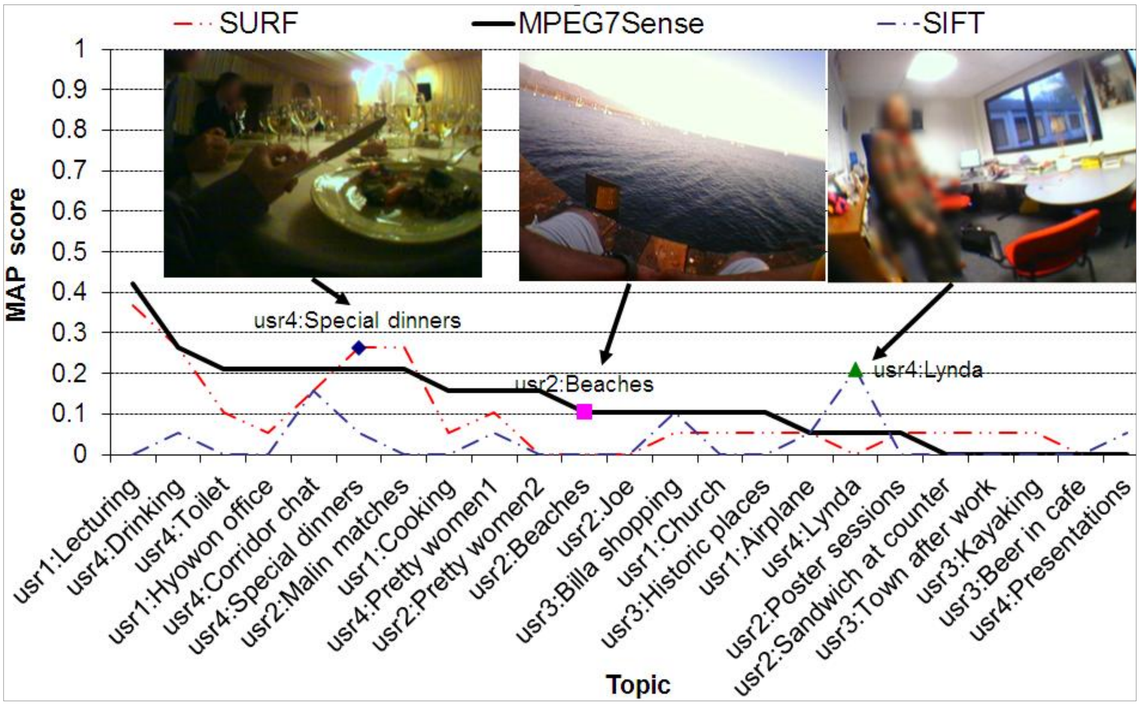

**Figure 6. Identifying events that are likely to be more distinctly encoded in the autobiographical memory system**

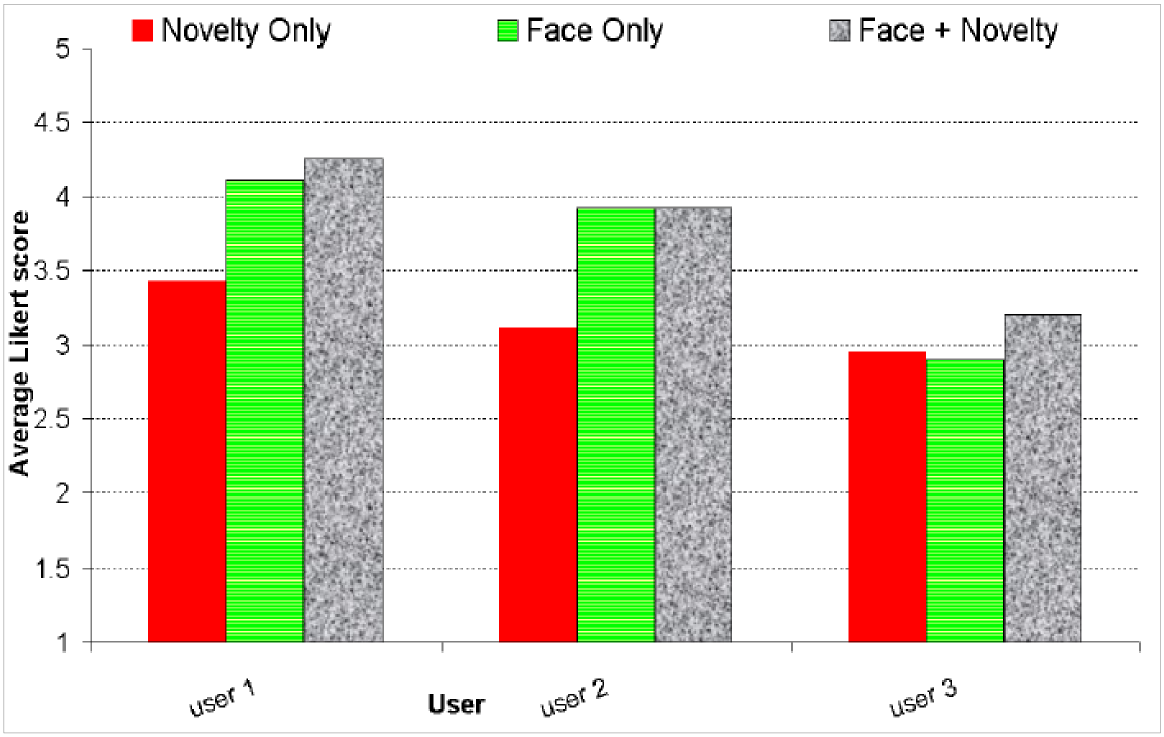

**Figure 7. Visual lifelogging "multi-axes" browser developed in 2010. The primary design goal was to provide multi-faceted retrieval of events to support person recollection. This browser aims to support search on the "who", "what", "when" and "where" axes of retrieval**