

Architecture and Challenges of Maintaining a Large-scale, Context-aware Human Digital Memory

Cathal Gurrin, Daragh Byrne, Noel O'Connor, Gareth J.F. Jones, Alan F. Smeaton

Adaptive Information Cluster, Dublin City University, Dublin, Ireland

Keywords: Photo, HDM, Memory, Lifelogging, Context Analysis, Content Analysis, Semantic Enrichment

Abstract *Recent advances in digital multimedia capture, transmission and storage has led us close to realising the MEMEX vision, in which much of an individual's personal information can be stored for retrieval at a later stage. These personal archives (Human Digital Memories (HDMs)) are constructed from a wide range of data sources across various media types. In this paper we examine the challenges of building, organising, browsing, searching and recommending from large (multi-million item per annum) visual HDMs. We present our findings based on an extensive period of HDM generation (2 years and 2 million images), illustrating the challenges of and architectural requirements for managing such archives.*

1 Introduction

In 1945, Vannevar Bush [3] envisioned a memex or "device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility." Bush's Memex vision would see a world where digital memory can supplement the shortcomings and fallibilities of human memory, where memories, experiences and life content such as that stored in diaries, data and encyclopedias can be accessed from one central repository with ease and efficiency. This idea is certainly not a revolutionary. Throughout history we can identify human achievements that have resulted from the desire to store memories, pass on life experiences and circumvent the limitations of human memory. From cave paintings, the invention of writing, to the WWW of today, many technological advances have allowed us to rely less on the fallible nature of human memory and more on an externally recorded surrogate of this memory. The concept of a MEMEX (or a Human Digital Memory) is another step (albeit a significant one) towards the goal of augmenting our unreliable memory with a surrogate.

In recent years, MEMEX has inspired the creation of tools such as MyLifeBits [9] and SLife [21], which enable the automatic capture of desktop computer content and activity. Gordon Bell of Microsoft has collected many years worth of personal and electronic data into one central repository using MyLifeBits. Additionally, Microsoft Research Cambridge

has built a prototype device, known as the SenseCam, designed to capture life experiences as a series of photographs [8].

Based on our experience of HDM development, and in realising part of the MEMEX vision, it is clear that a HDM should provide timely, precise and convenient access to an individual's HDM. While this goal may appear straightforward a HDM can be expected to contain an overwhelming amount of data which is constantly growing. For example, the SenseCam (a passive capture camera used in this research) can amass about 4,500 images in a typical day, each of which must be downloaded from the device, processed and indexed appropriately. Other sources of context data, such as a Bluetooth based social context logger can generate even greater volumes of data in very short periods. In addition, the HDM may contain every email, text message, web page that the user reviews, thus becoming a heterogeneous multimedia repository, expanding daily. This poses significant challenges both with regard to scalability, search and presentation of HDM content.

We outline, in this paper the challenges posed by HDM maintenance, and propose an example system architecture, which will allow a user to achieve just this by providing ubiquitous access to, and retrieval from, their lifelog data. We mainly consider accessing and supplementing visual lifelog data created using a SenseCam and the use of content and contextual processing to supplement these photographs. The architecture and challenges presented here are based on our experiences of managing large scale visual HDMs.

2 Related Research

Recording of personal life experiences through digital technology is a phenomenon we are increasingly familiar with: music players, such as iTunes, remembers the music we listen to frequently; our web activity is recorded in web browsers' History feature; and we capture important moments in our lifetime through photos and video. This concept of digitally capturing our memories is known as "lifelogging" and the storage device is the HDM. To enable visual lifelogging and memory capture, Microsoft Research in Cambridge, UK, have developed a device known as the SenseCam. The SenseCam is a small wearable device that passively captures a person's day-to-day activities as a series of photographs [8][9].



Figure 2. The Microsoft SenseCam (*Inset: as worn by a user*)

It is typically worn around the neck and, so is oriented towards the majority of activities which the user is engaged in. Anything in the view of the wearer can be captured by the SenseCam. At a minimum the SenseCam will automatically take a new image approximately every 50 seconds, but sudden changes in the environment of the wearer, detected by onboard sensors, can trigger more frequent photo capture. The device incorporates on-board sensors to detect changes in light levels, motion and ambient temperature and then determine when is appropriate to take a photo. Wearing a SenseCam, a wearer can very quickly build large and rich photo collections. Within just one week, over 20,000 images may be captured and over a year the lifelog photoset could grow to over one million images. The benefits of this include: the ability for a user to record events without having to sacrifice their participation, aiding memory and recall; and providing insight into a person's life and activities [13].

The rate at which a SenseCam photo collection can grow, presents a significant challenge to both browsing and retrieval, when compared with traditional photosets which typically contain several thousand images at most. Additionally, our days are composed of a series of events and encounters, so it is sensible to segment a day's activities into a set of discrete events [6]. Doherty [7] has investigated the optimal combination of the five onboard SenseCam sensors to define an event's uniqueness, suggesting that three sources work well: low-level image features (content), light level sensor (context), and accelerometer/motion (context). The segmentation process also allows us to calculate a measure of each event's uniqueness within the collection, enabling us to identify the more interesting events from a HDM.

An important component of our architecture leverages evidence from contextual cues from the surrounding environment of the user in order to enable better retrieval performance. The work of Byrne et al. [5] demonstrates that Bluetooth can be used to effectively enable people-based search and GPS can be employed to provide location-based retrieval of events within a SenseCam lifelog photoset.

3 Data Collection for the HDM

Capturing as much information (many sources) as possible actually works in favour of the HDM, as is the approach taken by the MyLifeBits [9] project at Microsoft. Integrating many sources increases the richness and the depth of life experience information which can be mined, thereby providing context to semantically enrich existing content.

Within our work one user captured data using a Microsoft SenseCam for a period of almost two years, resulting in a visual collection of two million photos. These photos were aligned with: the output of the onboard SenseCam device (accelerometer, temperature, light); a complete GPS log of all the individual's movements during this time; and sporadic periods of Bluetooth context-logging which enables the identification of individuals near the collection owner. This results in the daily collection of about 4,500 images, 1,000 Bluetooth interactions, 2,700 GPS points (when moving) and 57,600 sensor readings for each sensor.

While within our work to date we have not captured other sources of data there is the potential to collect a range of activity information or biometric data. Activity information can be extracted from both desktop (e.g. web pages visited, emails read, documents written) and mobile devices (e.g. SMS messages, log of phone calls, emails). Biometric information may also be captured from a wearer in order to determine physiological responses during their day-to-day activities and provide cues as to moments of significance. Potential sources of this data include: Galvanic Skin Response (GSR), Heat Flux, Skin Temperature and movement heart rate (i.e. excitement level).

3.1 Data Upload

The data within a HDM may be collected by a range of devices or computing environments and as such it will be necessary to bring these sources together into a central repository in which content processing and indexing can occur. Ease of upload is essential to reduce user overhead. Currently all our sources require the connection of the device to a computer using a USB (or similar) cable and a daily upload process that takes upwards of fifteen minutes to perform. Ideally however, minimal intervention from the user should be required. We envisage that the range of HDM collection devices should intelligently upload information. In the ideal scenario, the devices periodically attempt to connect to the repository either directly if in range of it or via a wireless internet connection and upload data as the opportunity arises and thereby minimize the user overhead.

3.2 Analysing Lifelog Visual Data

Based on a previous analysis we carried out on a million photo visual HDM [11], gathered also from a SenseCam, we conclude that the photos captured in over a year of continuous lifelogging are vastly different in visual content to conventional digital photos. We have found that these photos often don't have salient objects, many are either low quality or redundant (40%), and the types of scenes/objects captured differ greatly from conventional photo collections. This suggests that simply applying pre-existing concept detectors used for conventional digital photos to visual lifelog data will not be very effective. Instead a more specific set of considered concept detectors (such as driving, eating, talking or working) should be developed to address the unique challenges of organising a HDM. Additionally, this study reveals that many concepts found in HDMs relate to mundane

Concept	Present
Vehicles (external view)	0.93%
Road	1.95%
Inside of vehicle, not driving (airplane, car, bus)	4.33%
Indoors	37.39%
Door	1.12%
Outdoors	6.37%
Buildings	3.58%
Tree	1.85%
Grass	0.86%
Sky	2.93%
Vegetation	1.50%
Screen (computer/laptop)	20.51%
Newspaper/Book (reading)	1.79%
Meeting	4.63%
Office	15.06%
Food (eating)	4.15%
Hands	20.60%
Holding a cup/glass	0.81%
Faces	5.84%
People	12.83%

Table 1. Important Visual Concepts for HDMs

events that can be hidden unless specifically requested. Therefore automatic down-weighting of ‘*stop-events*’ (events that occur so frequently so as to be of little importance, such as driving or working at a desk) could be very useful within retrieval tasks. A listing of the twenty important concepts and their frequency of occurrence in a random selection of 51,697 SenseCam photos from five users are shown in Table 1.

4 Content Processing and Semantic Extraction

To support user search and content recommendation from a HDM and to bridge the semantic gap between raw HDM content and user expectations, it is necessary to apply a number of key processes to automatically enrich the semantics of the raw HDM data. Exactly what this semantic enrichment process is remains an open research question, but thus far, we have developed the following:

- Event segmentation to divide a HDM into discrete ‘events’ or temporally related activities.
- Visual Feature Extraction to add semantic annotations to the content.
- The integration of additional (context) sources of evidence for semantic enrichment, such as Bluetooth interactions, GPS and sensor output.
- Support for human annotation

We now examine each type of collection enrichment.

4.1 Detecting Events within SenseCam Images

Previous work [6],[7] recognised the need to automatically divide lifelog photosets into discreet events, where each event is represented by one or more key photos, instead of being represented as a sequence of hundreds or more individual photos. This challenge of event segmentation is quite similar to that of scene boundary detection in video as events or activities have an inherent underlying semantic meaning. In previous work we carried out an extensive evaluation [7] to

optimise event segmentation for lifelog images on a dataset of 271,163 images from five distinct users, who had manually groundtruthed 2,986 event boundaries. Recall of events was measured at 62.17% while 62.57% of all boundaries proposed by the system were valid boundaries (precision) when the automatic segmentation approach was compared to the manual groundtruth. This event segmentation approach attempts to identify periods of visual or sensory change, and identifies those occasions as most likely to be boundaries between distinct events or activities. Sensor data from the onboard SenseCam sources is employed along with MPEG-7 descriptors (colour layout, colour structure, scalable colour, and edge histogram) in order to automatically segment a day of images into distinct events.

4.2 Visual Feature Extraction

Given that a large proportion of the output from the SenseCam (and as a consequence, our HDMs as well) is visual in nature, visual feature extraction provides an excellent opportunity to support automatic concept detection and semantic enrichment of a HDM. Concept detection is used frequently in digital video retrieval to extract semantic concepts from frames of digital video footage [22] or in digital image retrieval. By matching the visual features of a frame within the footage to the properties of known ‘concepts’ (such as indoors, outdoors, people, crowd, etc.) the probability of the concepts occurrence in the frame is determined. Similar detection could be applied to the visual contents of a lifelog. Within a lifelog generic concepts such as those currently used in the domain of video retrieval could be employed but we suggest based on the analysis provided in [11] that a set of concepts tailored specifically to the contents of a lifelog should be considered. The detection of such concepts would allow us to semantically enrich an individual event within a HDM with automatic annotations. With the probability of a concepts presence within an event being known, the following would be enabled:

- Search of the HDM by concept or concept groups
- Event Linkage based on concept co-occurrence
- Down-weighting or hiding of ‘*stop-events*’

In order to enable concept detection, a typical approach would be to identify a set of known positive and negative examples for each concept and train a machine learning technique (e.g. SVM) based on these examples. In training the machine learning technique, as large as possible a training set should be used in order to maximise the reliability of output. Additionally, it is important that the positive examples should represent a diverse range of examples for a given concept, and as such, should occur in different events. This is because concept examples from a single event would tend to be highly temporally consistent and consequently too visually consistent. Finally we would envisage an active learning approach to concept detection in which as a user reviews search results or events they may indicate non-annotated content to belong to a given concept, or may feedback on incorrectly annotated images. Providing such feedback would continuously improve and adjust the concept detection process, retraining the classifiers as necessary. A

HDM, being a user specific collection would likely require the facility for users to create custom concepts for inclusion in the semantic enrichment process.

4.3 Integrating Context

Semantic enrichment by means of context analysis is essential to aid the semantic enrichment process. This context information will help to bridge the semantic gap between computer and human interpretation of visual content, and provides further contextual cues for the semantic enrichment of content contained in a visual HDM. In our work we analysed three sources of context information; these are, the user's location, user's environment and the output of various sensor devices.

Employing **location context** would typically require logging locations visited using GPS; we can annotate the places at which events occur both at the event and image level. This has proven to be extremely powerful in conventional photo archives as it allows the search and linking of content by location. Previous research [10] has shown that location is the most powerful access methodology to locate photos from a large personal archive and the same will should hold through for users accessing HDM archives. In addition, previous research on organising personal photo archives on mobile devices has illustrated the power of location to drastically reduce the average search time for photos on mobile content [12], which is a likely access modality for HDMs. Semantically enrichment of the HDM can also rely on linking to external WWW content, and location context provides a very powerful seed query to get additional content to enrich and supplement that of the HDM.

Employing **social context** information allows us to identify the interpersonal relationships between the people and locations contained within the HDM collection. By understanding this information we can seek to classify the relationships between those people and better understand the day-to-day activities they were involved in. It also allows us to better search, navigate and relate content contained within the HDM by the social factors which define or bind them. We envisage social context information would be gathered through the recording of Bluetooth encounters (people) and GPS location information (places). Social context can be simply described an understanding of the *places* in which and *people* that an individual interacts with and since a HDM contains the life experience of an individual this social context is very important for organising the HDM. One technique for detecting people automatically is by using Bluetooth-polling; people are rarely encountered in isolation and in their encounters quickly form regularly reoccurring highly cohesive groups. Using the Bluetooth information we can derive the people who have been encountered and the groups that they interact in. From this we can establish a social network in order to better understand the relationships between them. In [15], we extended Nicolai's work [19] in 'familiarity' which calculates the level of knownness between encountered individuals and the lifelog owner. This allows us to build a highly accurate map (or social network) of a

person's social interactions, which can then be used to group, search, navigate and recommend from their lifelog data.

Finally, **environmental context** can include sensor data from the SenseCam and from other sources such as biometric recording devices may allow additionally context about the lifelog owner to be captured; for example the accelerometer may distinguish between motion-based activities (running or walking) and the thermometer may distinguish events based on temperature (a hot afternoon or a very cold morning).

4.4 Supporting Human Annotation

While semantic concept detectors generally provide reliable output they are wholly dependent on the training examples provided. As such it is possible that there may be inaccuracies and incorrect or missed classifications within the collections. Support for a human intervention in the loop is therefore required. Within a HDM, the content owner should be able to verify and as necessary repair the annotations automatically produced by the system. They should also be able to provide their own custom annotations to any event or content item. An ideal approach to this would be the provision of text-based descriptions or, as in Flickr content, user-defined tags or keywords, which identify the content as belonging to a given concept. The provision of such tags may be a cue of the system to automatically generate a new classifier for future additions to the HDM. However, we cannot rely on a user annotating content with any regularity (it is unlikely to occur daily) or with broad coverage (given not likely for all content in a HDM), and as such user provided annotations should be considered a secondary source of evidence and not relied upon. From time to time, when a user does annotate content of personal importance, then the HDM should make it as easy as possible (most probably via spoken word) so as to minimise user overhead.

5 Indexing and Retrieval

Based on our experiences, and in order to be really useful, the HDM needs to be a silent repository (just like a human memory), easy to access and always available with instant access to just the correct information. This suggests that the indexing process needs to apply minimum cognitive load on the user, needs to apply semantic enrichment techniques to the data, and that retrieval needs to be context sensitive, fast and accurate. This poses major challenges, for both indexing and retrieval. For indexing the challenges can be thus summarised:

- To minimise the cognitive load on the user when uploading and indexing the content.
- To automatically organise the content into events which are meaningful units to the user.
- To automatically apply semantic enrichment techniques to the events and photos within the HDM, so as to provide easy and powerful retrieval.

Minimising the cognitive load on the user suggests that manual annotation of memory items (events) needs to be an automatic process, relying on human interactions only for verificitive processes and only when needed. Event detection

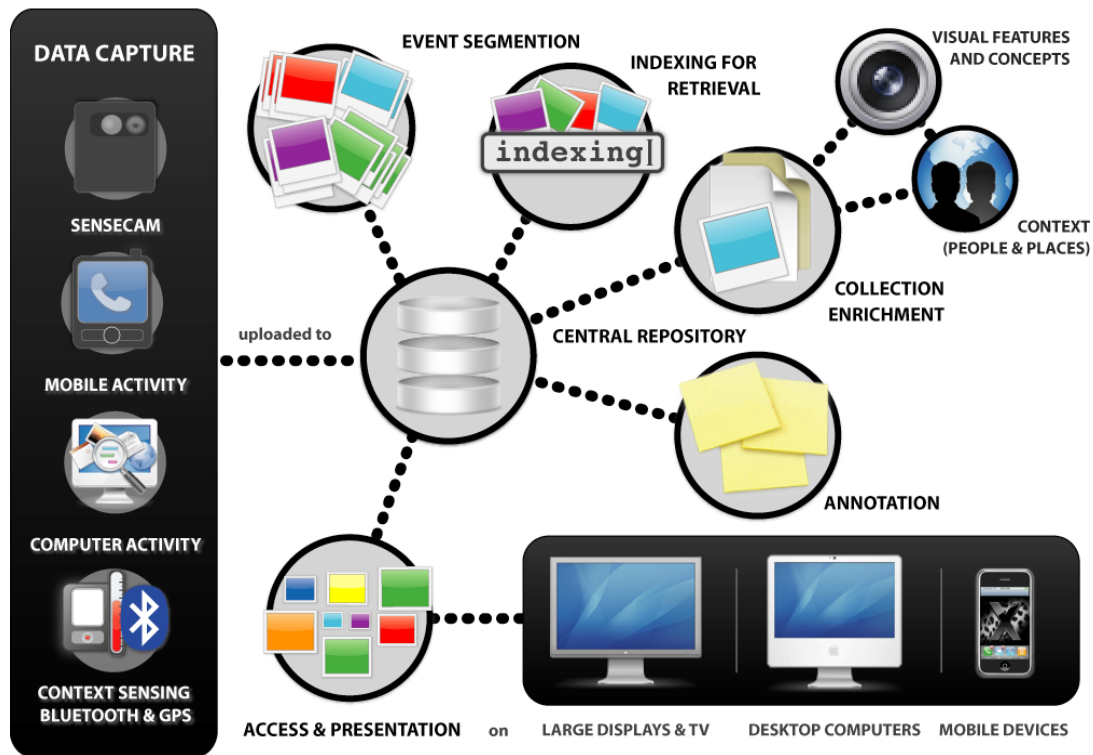


Figure 2. Overview of the proposed architecture

as a requirement is necessary as it helps to organise the HDM into events, which are a reasonable unit of retrieval. The application of semantic enrichment techniques (context and content) to the events is also a key challenge, and we have identified some of the unique visual concepts earlier, which differ from concepts required for conventional digital photos. For retrieval from HDMs, the challenges are:

- To provide fast search facilities for millions of photos per year, many tens of millions of sensor data, and tens of thousands of human interactions.
- To automatically link between related events and provide a browsing/search mechanism.
- To provide extreme precision of retrieval.
- To support context sensitive retrieval, in that the HDM can seek out and locate events based on the current actions of the user.
- To be always available to the user.

The retrieval challenges posed are for the most part research problems yet to be solved, though we believe that they are achievable with sufficient quality of indexing. The multi-modal nature of HDM data provides, both a unique challenge, and at the same time, a unique opportunity to achieve these retrieval challenges and provide for high precision and highly efficient indexing and retrieval of HDM data.

6 Presentation Challenges

The core objective of a HDM is to make life experience data accessible to its owner in order to successfully supplement the deficiencies of our own human memory. Consequently, careful consideration of the modes, methods and situations of access to this information is required.

Simply organising a HDM by date or time and presenting the user with hundreds or thousands of photos is not a realistic scenario. Likewise, playing a stop-motion video of a day for a user is likely to take in the order of 4-5 minutes to watch is not likely to satisfy a user's information need. Some of the presentation challenges have already been explored. The 'wearable remembrance agent' [20] seeks to augment a user's real-world interactions with digital content such as documents. Both Byrne [4] and Lee [16] have already considered novel structuring and presentation of visual lifelog content, while the MemoryLane application [18] explores the use of digital memory access for the elderly on PDAs. Other issues in presentation of content such as context-aware retrieval of information [2],[14] also need to be examined. Bush [3] visualised his MEMEX as an environment where the owner could extract and explore their life data. While Bush's MEMEX vision presented a desktop approach to HDM access, we firmly feel that a true MEMEX should not be bound to just a desktop but rather, should be additionally available on mobile devices. In fact, the real value of such life data might only be realised in situations where the owner is on the move and away from a desktop computer. We suggest three core modalities for access:

- Desktop access using a conventional computer which would provide for powerful access to lifelog contents.
- Mobile access to support anytime access to the HDM content, as and when a user requires; should be designed for life experience access on the move, and should support context-aware retrieval.
- Semi-public or shared display access whereby a user could access HDM content with friends and family, and integrate multi-user context into the

query facility, thereby sharing life content on large situated displays with others. Scenarios of use would be sharing an important event, such as a trip, with family on the living room television or perhaps placing content of interest on a shared public display within an office environment for colleagues to review.

There is also the possibility that in future life content may be more widely shared through for example social networking sites or as part of weblog posts.

7 HDM Architecture & Conclusions

In this paper we have presented the challenges of automatically maintaining a large multi-modal HDM; in addition we have presented an overview architecture for a large context aware HDM based on our experiences of working with large HDM archives. This overview architecture, presented in Figure 2, illustrates the major components of the HDM from data capture through to processing and eventual presentation to the user. Implementation issues such as index structures, data transfer methods and interface requirements are not illustrated or discussed in this paper.

In summary, the requirements for a useful HDM, based on our experiences, can be thus summarised:

- The need for fast and efficient access on a variety of presentation devices.
- A requirement for segmentation into meaningful units such as events.
- A requirement for automatic semantic enrichment by utilising context and content analysis.
- Support for ubiquitous user annotation of content.
- A requirement for high precision and context sensitive retrieval, along with event linkage within the HDM.

The architectural specification and challenges presented are designed for a single-user HDM (does not address multi-user HDMs or HDM sharing). In future work, we will examine multi-user HDMs and the resulting issues for HDM sharing. Many of the presentation considerations remain to be explored particularly multi-platform access to lifelog content.

Acknowledgements

We acknowledge and thank Microsoft for their support of our lifelogging research. This work was also supported by the Irish Research Council for Science Engineering and Technology and by Science Foundation Ireland under grant 03/IN.3/I361 and was partially supported by the COST292 Action on Semantic Multimodal Analysis of Digital Media.

References

[1] The AceMedia project. <http://www.acemedia.org>.
 [2] P.J. Brown, G.J.F. Jones. "Exploiting contextual change in context-aware retrieval", *Proceedings of the 2002 ACM symposium on Applied Computing*, Madrid, Spain, pp. 650 - 656, (2002).

[3] V. Bush. "As We May Think", *In The Atlantic Monthly*, 176, 1, pp. 101-108, (July 1945).
 [4] D. Byrne "SenseCam Flow Visualisation for LifeLog Image Browsing", *BCS IRSG Informer No. 22*, pp. 4-7, (2007).
 [5] D. Byrne, B. Lavelle, A.R. Doherty, G.J.F. Jones, A.F. Smeaton. "Using Bluetooth & GPS Metadata to Measure Event Similarity in SenseCam Images", *In Proceedings of the Fifth International Conference on Intelligent Multimedia and Ambient Intelligence*, pp. 1454-1460 (July 2007).
 [6] A.R. Doherty, A.F. Smeaton. "Automatically segmenting lifelog data into events". *In Proceedings of 9th International Workshop on Image Analysis for Multimedia Interactive Services*, (2008).
 [7] A.R. Doherty, A.F. Smeaton, K. Lee, D. P. Ellis. "Multimodal segmentation of lifelog data", *In Proceedings of RIAO 2007*, (2007).
 [8] J. Gemmell, L. Williams, K. Wood, R. Lueder, G. Bell. "Passive Capture and Ensuing Issues for a Personal Lifetime Store", *In Proceedings of the 1st ACM workshop on Continuous archival and retrieval of personal experiences*, pp. 48-55, (October 2004).
 [9] J. Gemmell, A. Aris, R. Lueder. "Telling Stories with MyLifeBits", *In Proceeding of IEEE International Conference on Multimedia and Expo (ICME 2005)*, Amsterdam, Netherlands, (July 2005).
 [10] A. Graham, H. Garcia-Molina, A. Paepcke, T. Winograd, "Time as essence for photo browsing through personal digital libraries", *In JCDL 2002*. Portland, Oregon, USA, (2002).
 [11] C. Gurrin, A.F. Smeaton, D. Byrne, N. O'Hare, G.J.F. Jones, N. O'Connor. "An Examination of a Large Visual Lifelog". *In AIRS 2008*, Harbin, China, (January 2008).
 [12] C. Gurrin, G.J.F. Jones, H. Lee, N. O'Hare, A.F. Smeaton, N. Murphy. "Mobile Access to Personal Digital Photograph Archives", *In 7th International Conference on Human Computer Interaction with Mobile Devices and Services*, Salzburg, Austria, (September 2005).
 [13] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, K. Wood. "SenseCam: A Retrospective Memory Aid", *In Proceedings of the 8th International Conference on Ubicomp 2006*, pp. 177 - 193, (September 2006).
 [14] G.J.F. Jones. "Challenges and Opportunities of Context-Aware Information Access", *International Workshop on Ubiquitous Data Management*, pp. 53-62, (2005).
 [15] B. Lavelle, D. Byrne, C. Gurrin, A.F. Smeaton, G. Jones. "Bluetooth Familiarity: Methods of Calculation, Applications and Limitations", *In MIRW 2007*, (September 2007).
 [16] H. Lee, A.F. Smeaton, N. O'Connor, G.J.F. Jones. "Adaptive Visual Summary of LifeLog Photos for Personal Information Management", *In Proceedings of AIR Workshop - 1st International Workshop on Adaptive Information Retrieval*, Glasgow, Scotland, (October 2006).
 [17] W.H. Lin, A. Hauptmann. "Structuring continuous video recordings of everyday life using time-constrained clustering", *SPIE Symposium on Electronic Imaging*, (2006).
 [18] S. McCarthy, H. Sayers, P. McKeivitt. "Investigating the Usability of PDAs with Ageing Users". *In Proceedings of the 21st BCS HCI Group Conference*, Lancaster, U.K., (September 2007).
 [19] T. Nicolai, E. Yoneki, N. Behrens, H. Kenn. "Exploring Social Context with the Wireless Rope". *1st International Workshop on MOBILE and Networking Technologies for social applications (MONET'06)*, Montpellier, France (2006).
 [20] B.J. Rhodes. "The wearable remembrance agent: A system for augmented memory", *In Proceedings of The First International Symposium on Wearable Computers (ISWC '97)*, Cambridge, Mass, pp. 123-128, (October 1997).
 [21] Slife. Retrieved from <http://www.slifelabs.com/>
 [22] C.G.M. Snoek, B. Huurnink, L. Hollink, M. de Rijke, G. Schreiber, M. Worring. "Adding semantics to detectors for video retrieval", *IEEE Transactions on Multimedia*, 9(5):975-986, (August 2007).